Structural basis of porcine RNase 4 recognition

Liang, Shutian

Award date:
2015

Awarding institution:
University of Bath

Structural basis of porcine RNase 4 recognition

Shutian Liang

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

October 2014

COPYRIGHT

Acknowledgements

First of all, I would like to thank my supervisor Professor Ravi Acharya for giving me the opportunity to work on this project, as well as all the help, guidance, and motivations you have given me during my study. I am grateful to the University of Bath Graduate School for funding my PhD study.

I would like to thank the station scientists at Diamond Light Source, Didcot, Oxfordshire, UK, for their support and help during X-ray data collection.

I am thankful to all my colleagues from the Lab 0.34, both past and present, for their support and help during my study. Special thanks go to Drs Geoffrey Masuyer, Shalini Iyer, and Saima Rehman for sharing their knowledge on X-ray crystallography, and Charlotte Harrison for the proofreading.

I would like to extend my thanks to all my friends, both in China and in the UK, for all the care and support they have given me.

Last but not least, I would like to express my deepest and sincere appreciations to my family and relatives, especially my parents, to whom this thesis is dedicated. It was your understanding and encouragement that kept me moving forward and finally completed my PhD study and this thesis.

在此，我要感谢我的父母，给予我生命并竭尽全力给予我接受教育的机会。是你们多年来不断给予我支持和鼓励，帮助我取得了今天的成绩。养育之恩，无以回报，你们永远健康快乐是我最大的心愿；你们的无限理解和鼓励支持着我实现着自己的梦想。谨以此文献给我挚爱的亲人们！

Abstract

Bovine pancreatic ribonuclease (RNase A) and its homologues are pyrimidine-specific ribonucleases widely present in mammals, birds, amphibians, reptiles, and fish. RNases recognise a specific sequence – an adenosine 3' to a pyrimidine – on RNA, and cleave the molecule on the 3' side of the 3'-phosphate of the pyrimidine base. Extensive studies have been carried out on the RNase A homologues, including eosinophil-derived neurotoxin (EDN; RNase 2), eosinophil cationic protein (ECP; RNase 3), and angiogenin (ANG; RNase 5), and revealed distinct biological functions: EDN and ECP are involved in neurotoxicity, and ANG possesses angiogenic activity. RNase 4, although being discovered for a long time, is not as well characterised as much as other RNases.

RNase 4 has been found in several mammalian species including a few primates, porcine, bovine, and rodents. The mature protein of RNase 4 consists of 119 amino acids, making it the shortest amongst all RNase A homologues. It has a higher inter-species similarity than its homologues, and such high evolutionary conservation suggests that RNase 4 has a more specialised function than RNA degradation. While RNase A, EDN, ECP, and ANG show cytidine preference, RNase 4 has a strong preference for uridine, which can be reversed back to cytidine by a single amino acid substitution of Asp-80, as shown by studies performed with porcine RNase 4 (also known as PL_3).

In this study, we used PL_3 as a model to study the substrate specificity of RNase 4, and have solved four structures, including PL_3, PL_3 D80A mutant, and these two proteins in complex with dUMP and dCMP respectively. PL_3 adopts the classic kidney-shaped RNase A fold, and residues forming the substrate binding subsites occupy similar positions as those in human RNase 4 and the prototypic RNase A. The structure of PL_3 D80A mutant resembles that of the wild type

PL_3, and only hydrogen bond interactions between the side chains of Asp-80 and Arg-101 are lost.

The structure of PL_3-dUMP complex revealed interactions between the dUMP and residues Arg-7, His-12, Thr-44 and Phe-117 of PL_3, which were also observed in the structure of human RNase 4 in complex with dUp. The additional hydrogen bonds identified between dUMP and residues Gln-11, Lys-40, Asn-43, and Lys-119 of PL_3, as well as the absence of the interactions between Arg-101 of PL_3 and the ligand that were present in the hRNase 4-dUp structure, could be due to the flexibility of the mononucleotide ligand.

The crystal structure of PL_3 D80A-dCMP complex presents a small number of hydrogen bond interactions between the protein and the dCMP ligand, which might be sufficient to stabilise the ligand in the B_1 subsite, as the repulsion force on the dCMP ligand from the side chain of Arg-101 is absent in the PL_3 D80A mutant. This is because in the D80A mutant, Ala-80 cannot provide hydrogen bonding that would hold the side chain of Arg-101 towards the B_1 subsite.

The activities of RNases can be inhibited by a 50 kDa cytosolic protein, the natural ribonuclease inhibitor (RI). RI binds to all the members of the RNase A superfamily, thus regulating the cytoplasmic RNA levels and protecting cells from inappropriately secreted RNases. The interactions between RNase and RI are tight, reversible, and in a 1:1 ratio. Several complex structures of RNase-RI from various species have been determined, and the residues in the interfaces between RNase and RI are conserved in all of the complexes. Studies revealed a 17-fold tighter interaction between PL_3 and human RI than RNase A, making it very interesting to study the structure of the PL_3-RI complex and characterise the interactions between PL_3 and RI proteins.

To date, we have established purification protocols for both proteins, and the next step towards the structure of PL_3-RI would be to prepare and purify the protein complex, subject the protein complex to crystallisation experiments, and eventually lead us to the structural determination of PL_3-RI complex.

Table of contents

List of figures

List of Tables

List of abbreviations

2'-dCpA	2'-deoxycytidylyl-3',5'-adenosine
2'-dUpA	2'-deoxyuridylyl-3',5'-adenosine
2'-F-dCpA	2'-deoxy-2'-fluoro cytidylyl-3',5'-adenosine
2'-F-dUpA	2'-deoxy-2'-fluoro uridylyl-3',5'-adenosine
Å	angstrom
ANG	angiogenin
ALS	amyotrophic lateral sclerosis
ASU	asymmetric unit
BSA	bovine serum albumin
CpA	cytidylyl-3',5'-adenosine
Da	dalton
dCMP	deoxycytidine 5'-monophosphate
DLS	Diamond Light Source
DNA	deoxyribonucleic acid
DTT	dithiothreitol
dUMP	deoxyuridine 5'-monophosphate
dUp	deoxyuridine phosphate / 2'-deoxyuridine 3'-monophosphate
EAR	eosinophil-associated ribonuclease
E. coli	*Escherichia coli*
ECP	eosinophil cationic protein
EDN	eosinophil derived neurotoxin
EDTA	ethylene-diamine-tetraacetic acid
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
ExPASy	Expert Protein Analysis System server
GST	glutathione S-transferase
hANG	human angiogenin
hECP	human eosinophil cationic protein
hEDN	human eosinophil derived neurotoxin
HIV	human immunodeficiency virus

HPLC	high performance liquid chromatography
hRI	human ribonuclease inhibitor
hRNase 4	human ribonuclease 4
IPTG	isopropyl-β-D-1-thiogalactopyranoside
IR	isomorphous replacement
kDa	kilodalton
LB	Luria-Bertani
LIC	ligation independent cloning
LRRs	leucine rich repeats
MAD	multiple-wavelength anomalous diffraction
MBP	maltose-binding protein
MR	molecular replacement
mRNA	messenger ribonucleic acid
MPD	methyl-2,4-pentanediol
NCBI	National Centre for Biotechnology Information
OD_{260}	optical density at 260 nm
OD_{600}	optical density at 600 nm
ONC	onconase
PCR	polymerase chain reaction
PCT	pre-crystallisation test
PDB	Protein Data Bank
PEG	polyethylene glycol
PL_3	porcine liver ribonuclease 4
pRI	porcine ribonuclease inhibitor
PTM	post-translational modification
RCSB	Research Collaboratory for Structural Bioinformatics
RI	ribonuclease inhibitor
RNA	ribonucleic acid
RNase	ribonuclease
RNase A	ribonuclease A
RNase 4	ribonuclease 4
rRNA	ribosomal ribonucleic acid
RSV	respiratory syncytial virus

SAD	single-wavelength anomalous diffraction
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SNR	signal-to-noise ratio
TB	Terrific Broth
TFA	trifluoroacetic acid
tRNA	transfer ribonucleic acid
UniProtKB	Universal Protein Knowledgebase
UpA	uridylyl-3',5'-adenosine
UV	ultraviolet
UVC	uridine 2',3'-cyclic vanadate
v/v	volume to volume
wtPL_3	wild type porcine liver ribonuclease 4
w/v	weight to volume

Chapter 1. Introduction to protein X-ray crystallography

1.1 Introduction

X-ray crystallography refers to the experimental science of the determination and analysis of three-dimensional structures of molecules. Max von Laue in 1912, William L. Bragg and William H. Bragg in 1913 first shed light on X-ray diffractions of crystals (Bragg, 1914), and were recognized by Nobel Prizes in Physics in 1914 and 1915. A century later, to commemorate the enormous successes of crystallographic researches that had been achieved in almost every branch of science, the United Nations announced 2014 to be "the International Year of Crystallography".

The importance of X-ray crystallography in the area of Biochemistry is best reflected by its impact on the determination of macromolecule structures. After the intensive studies done on genomics, proteomics has become the next step in the study of biological systems. Proteins are the fundamental macromolecules of an organism and play crucial roles in biological processes. Interactions between proteins and other biomolecules have been identified in many cellular events, to name a few: cell growth and proliferation, differentiation and migration, structural and mechanical support, host immune defence, energy and nutrients storage, and enzymatic catalysis. The functions of proteins are defined by their physical features, thus the determination of protein structures is important and necessary for researchers to understand biological processes at the molecular level.

In protein crystallography, the protein sample can either be obtained from other manufacturers, or be produced in one's own laboratory. The production of protein often involves molecular cloning of the gene of interest, protein expression, and purification. The protein sample is crystallised and used in the collection of X-ray diffraction data of the crystals. The diffraction data are processed with the aid of computational programmes from where the three-dimensional structure of a protein may be solved (Figure 1.1). This chapter will

discuss in general the methodologies involved in the determination of protein structures.

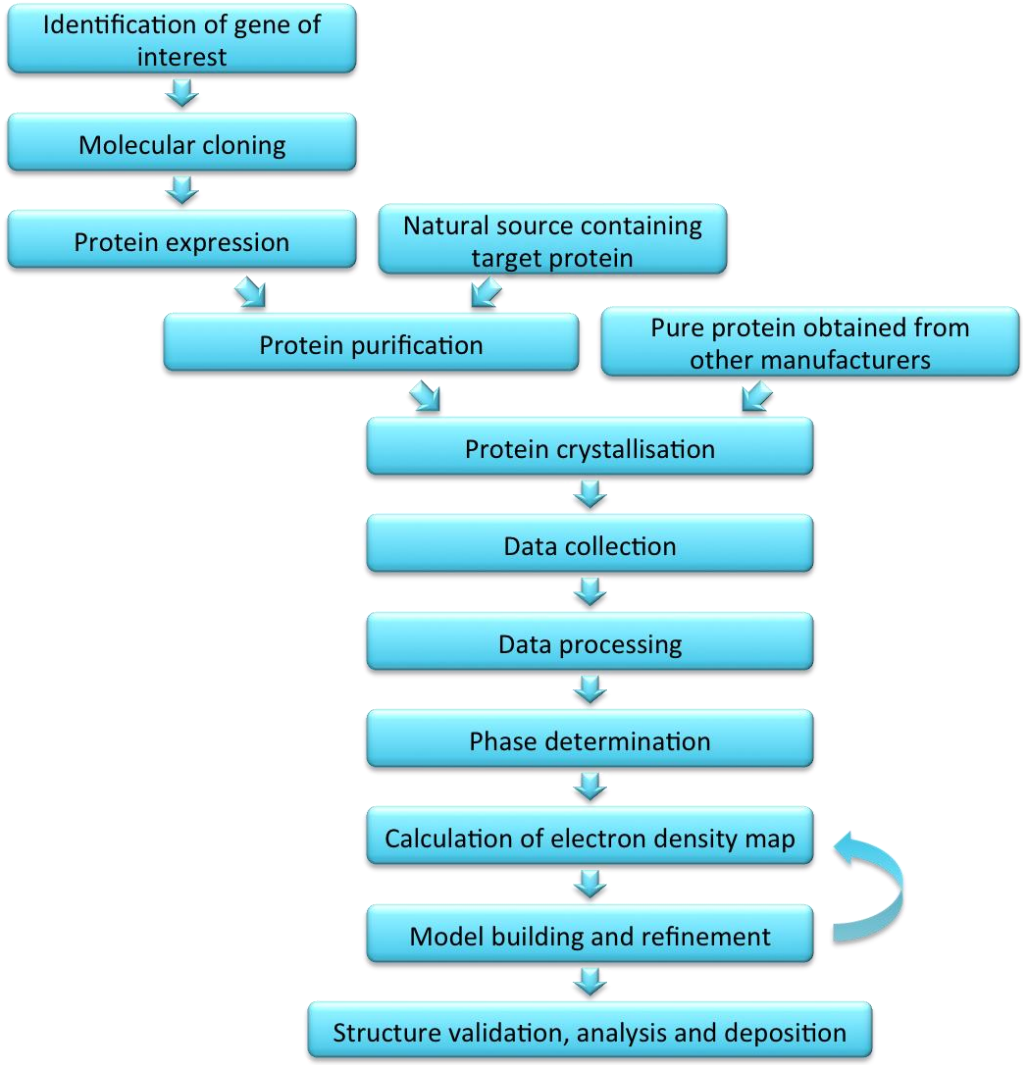


Figure 1.1. Steps involved in protein structure determination using X-ray crystallography.

1.2 Molecular cloning

Molecular cloning refers to the experimental methods that researchers use to isolate the DNA sequence of interest from an organism, combine it into a small DNA molecule also known as a vector, and transfer the recombinant DNA molecule into a host cell where the DNA of interest can be expressed. Molecular cloning methods have a very wide application in modern biochemical and medical research including protein X-ray crystallography.

The first step to a successful construction of a protein expression system is the identification and isolation of the DNA sequence of target protein. By the use of bioinformatics databases, the DNA sequence of target protein can be obtained and analysed. First, the DNA sequence of the target protein needs to be isolated from the chromosomal DNA and amplified by Polymerase Chain Reaction (PCR). The PCR product is subsequently treated with restriction endonuclease enzymes to create sticky ends or blunt ends; such treated DNA fragments are also known as inserts. In parallel, a selected DNA vector, often with additional features to ease subsequent purifications (such as affinity tags) and/or aid a better solubility (such as MBP-tag and Trx-tag) (Malhotra, 2009), is treated with the same enzymes to create a linear vector with complementary ends to those of the insert. The insert and linear vector are ligated with the presence of DNA ligase to produce a recombinant DNA molecule. This most common method was developed in early 1970s (Jackson *et al.*, 1972; Cohen *et al.*, 1973; Lobban and Kaiser, 1973), and more creative cloning techniques to assemble recombinant DNA molecules have been developed through the years. One of such techniques, ligation-independent cloning (LIC), was used in this project.

In LIC cloning, no restriction endonucleases or DNA ligase are required, the creation of sticky ends is solely fulfilled by the 3' to 5' exonuclease activity of T4 DNA polymerase in the presence of dATP. Insert and linear vector with complementary sticky ends are mixed to allow annealing. The product – the annealed recombinant DNA plasmid – can be used in the subsequent transformation or transfection of host cells.

1.3 Protein expression

The recombinant DNA molecules need to be transfected into an appropriate host to aid protein expression. The choice of the host, also known as the expression system, depends on the properties of the recombinant protein.

If post-translational modifications (PTMs), such as glycosylation and phosphorylation, are required for the correct folding and activity of the recombinant protein, a eukaryotic expression system should be used. Yeast cells (for example *Saccharomyces cerevisiae*), insect cells, and mammalian cells are suitable candidates for this purpose. If these modifications are not required, a bacterial expression system will be the best choice, with *Escherichia coli* (*E. coli*) being the most commonly used amongst all.

Often, the gene of the protein is carried into *E. coli* cells by an expression vector with a T7 promoter, and the protein expression is induced by the addition of a synthetic analogue of lactose, isopropyl-β-D-1-thiogalactopyranoside (IPTG). IPTG is able to initiate the transcription of the T7 gene that is anchored in the genome of *E. coli*, and the T7 RNA polymerase produced will start the transcription of the expression vector, hence the target gene. The T7 RNA polymerase is so selective that almost all the resources of the cells will be utilised in the expression of the protein of interest, resulting in a high yield of the protein within a few hours of induction.

1.4 Protein purification

Protein purification refers to the process of isolating the target protein from a natural source or a recombinant expression system. In protein crystallography, impurities often impede crystal growth. Although there is no clear cut on the requirement for the purity of protein samples, ~95% purity, represented by a single band on a well-loaded SDS-PAGE gel, is often considered to be a good starting point for protein crystallisation (Rupp, 2010).

If the target protein is an extracellular protein, the tissue fluid or the supernatant of expression medium can be directly subjected to purification columns.

However, if the protein is expressed intracellularly, whether it is from a natural or a recombinant source, the first step in protein purification is to disrupt the cells and release the protein into solution. Cell disruption can be achieved either by physical force (namely mechanical disruption, liquid homogenisation, sonication, repeated freeze/thaw, and manual grinding) or the use of chemical cell lysis reagents. The choice of the technique mainly depends on the fragility and stability of the protein, and the amount of sample to be processed. If the protein of interest is soluble, it will be present in the supernatant after the centrifugation of the cell lysate; if the protein is in its insoluble form, also known as inclusion bodies, it will have to be solubilised and refolded prior to or during purification.

Protein purification usually involves one or a few chromatographic techniques. In general, the sample containing the target protein is flowed through a column packed with materials designed for a specific type of chromatography. The type of chromatography used is dependent on the properties of the target protein being exploited, and is chosen so that the target protein interacts with the column material differently from other macromolecules, which leads to its separation from the mixture. Commonly used chromatographic techniques and the properties they are based on are listed in Table 1.1. Most of the listed chromatography techniques were employed in this study and will be discussed briefly below.

Table 1.1 Chromatographic techniques commonly used in protein purification.

Property	Chromatographic techniques
Charge	Ion exchange chromatography
Ligand specificity	Affinity chromatography
Hydrophobicity	Hydrophobic interaction chromatography Reverse phase chromatography
Size	Size exclusion (also known as gel filtration) chromatography

Ion exchange chromatography separates proteins based on their net surface charges (Figure 1.2a). The net surface charge of a protein molecule varies

according to the surrounding pH. If the pH is higher than the isoelectric point (pI) of the protein, the protein is net negatively charged, and will bind to a positively charged column (anion exchange column); if the pH is lower than its pI, the protein is net positively charged, and will bind to a negatively charged cation exchange column. Bound protein can be eluted either by altering the surrounding pH, or by increasing the ionic strength of the system.

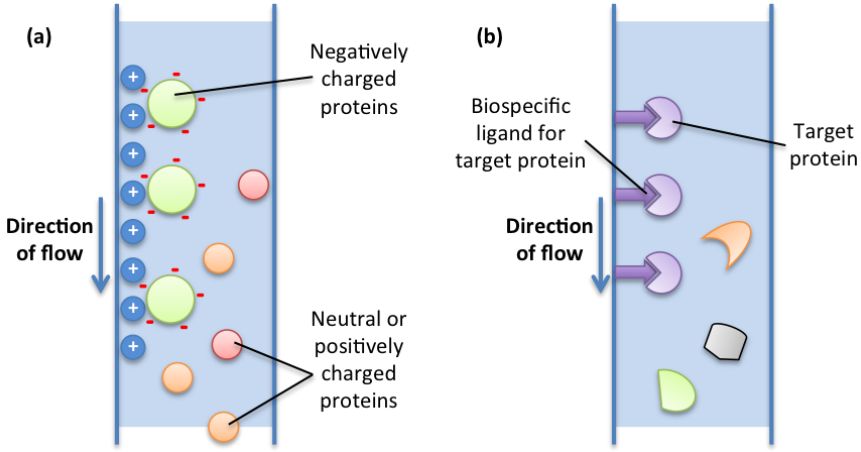


Figure 1.2. Separation principles of ion exchange chromatography and affinity chromatography. (a) In an anion exchange column, only negatively charged proteins bind to the positively charged column; (b) In an affinity column, only proteins with specific recognition towards the ligands bind to the column.

Affinity chromatography is a very popular protein purification method. It traps target protein by specific biointeractions (Figure 1.2b). As mentioned in Section 1.2, the chosen vector often adds additional features to the protein that enable the biointeractions to occur. These features are known as fusion tags, and they provide the recombinant protein with a specific ligand-binding affinity. Common tags include poly-histidine tag (His-tag), glutathione S-transferase tag (GST-tag), maltose binding protein tag (MBP-tag), and S-tag to name a few.

Once highly purified protein sample is obtained, the fusion tags can be cleaved off from the protein by a specific protease, whose cleavage site is engineered in between the tag and the protein (Wilchek and Chaiken, 2000). If the target

protein is known to have a high-affinity binding partner that can be used to coat the matrix of the affinity column, it can be expressed without tags. Since the recombinant protein is usually the only one in the crude sample with the ligand binding affinity, the separation of the target protein with high purity can be achieved in a single step.

Both hydrophobic interaction chromatography and reverse phase chromatography separate proteins according to their hydrophobicities (Figure 1.3). In hydrophobic interaction chromatography, protein samples are loaded onto a column in a buffer with high salt concentration. The high salt concentration drives the interactions between the hydrophobic proteins and the alkyl- derivatised stationary phase. By gradually lowering the salt concentration, proteins can be eluted in order of hydrophobicity.

Reverse phase chromatography utilises a more hydrophobic stationary phase, thus requiring organic solvents such as acetonitrile or methanol to elute the proteins. Most commonly, reverse phase chromatography is combined with high-performance liquid chromatography (HPLC), where a high pressure is applied to drive solutions move more quickly through the column to improve resolution (Regnier, 1983).

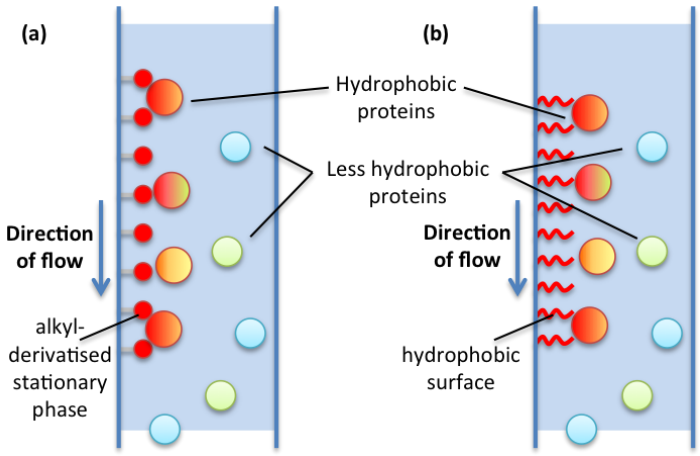


Figure 1.3. Separation principles of (a) hydrophobic interaction chromatography and (b) reverse phase chromatography.

Size exclusion chromatography, also known as gel filtration chromatography, separates protein samples based on their molecular weights (Porath and Flodin, 1959; Eisenstein, 2006) (Figure 1.4). The matrix of a size exclusion column is made of porous particles, to which proteins do not bind but travel through. Large proteins can only travel through the spaces between the particles thus elute first, whereas small proteins can penetrate every region of the matrix and elute late. One advantage of size exclusion chromatography is that the buffer composition can be varied to suit the requirements for further purifications or analysis without altering protein separation. Because of this, size exclusion chromatography is often considered to be the final polishing step of protein purification.

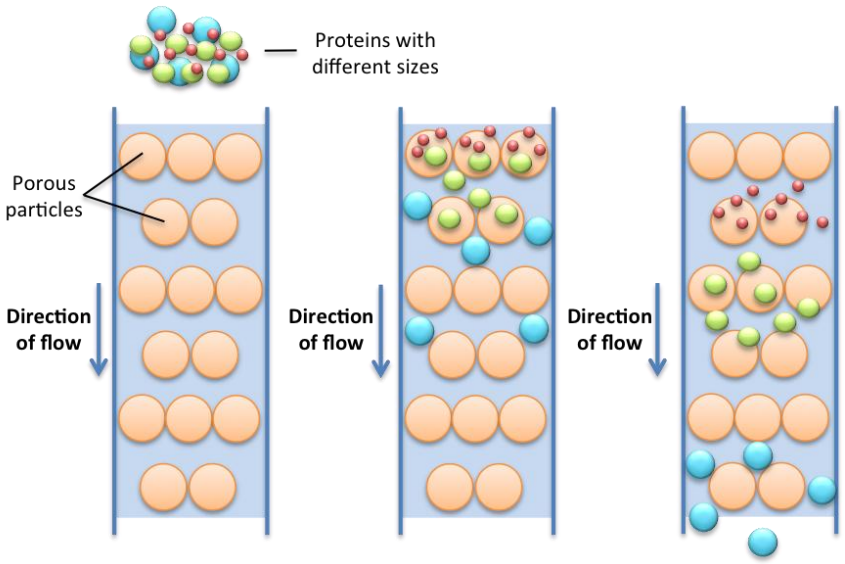


Figure 1.4. Separation principle of gel filtration chromatography.

The purity of a purified protein is analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), and its concentration can be determined by means of Bradford assay or absorbance of light at 280 nm. Biophysical (such as circular dichroism analysis, dynamic light scattering analysis, and mass spectrometry analysis) and biochemical characterisations of the protein should be performed to confirm the purified protein is mature, homogeneous and functionally active. If necessary, lyophilisation or

ultracentrifugation can be applied to the purified protein to obtain a more concentrated sample for subsequent protein crystallisation trials.

1.5 Protein crystallisation

In order to obtain diffraction data, the protein must be crystallised first. Crystallisation occurs when the concentration of the protein is greater than its solubility at that specific temperature, which is known as the supersaturation state. It is a non-equilibrium state and has a tendency to re-establish equilibrium by the formation and development of a solid state such as crystals (McPherson, 2004). In principle, two major events happen during protein crystallisation: nucleation of the crystal, and crystal growth. Nucleation is a process where protein molecules aggregate into ordered and stable clusters in the current environment that serve as the nuclei. Crystal growth is the subsequent aggregation of protein molecules to the nuclei, during which the solid state becomes more attractive than the solution state (Figure 1.5).

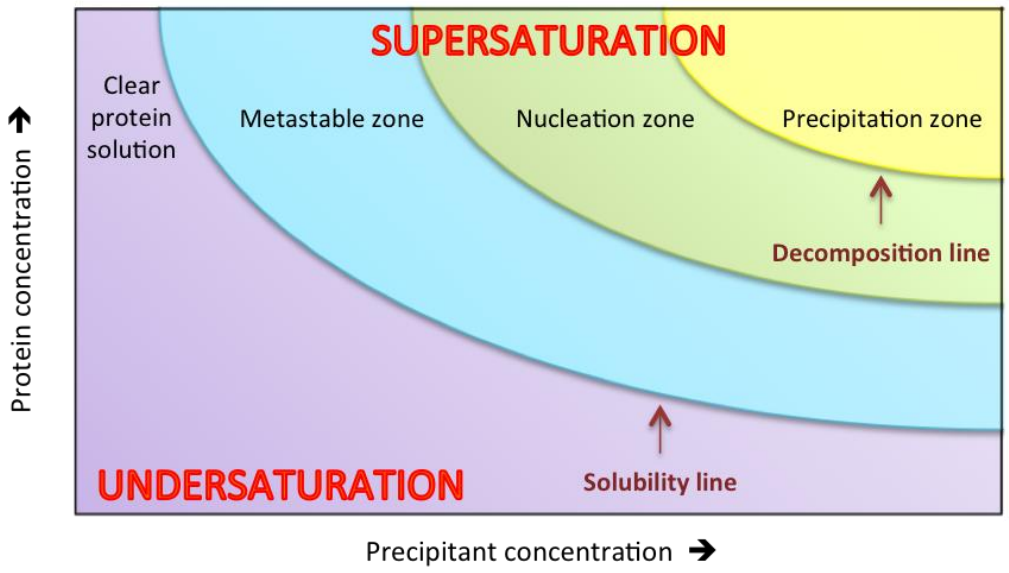


Figure 1.5. A phase diagram of crystallisation process. The precipitant represents the entire crystallisation cocktail.

Although the principle of protein crystallisation is similar to that of the crystallisation of small molecules, it is more complicated in practice. There are various factors that can affect crystal growth, including ① purity and concentration of protein, ② buffer systems, ③ pH, ④ types and concentrations of salts and precipitants, ⑤ organic solvents, ⑥ detergents, and ⑦temperatures. Despite the huge amount of research done on protein crystallisation, it is still impossible to predict the perfect crystallisation condition for a protein. Hundreds or even thousands of conditions with various combinations of the factors mentioned above may need to be screened to identify the preferred conditions for crystallisation of a protein. Such screenings are often performed using pre-made, commercially available screens and are set up using automated crystallisation robots. Once initial hits are obtained, more subtle variations of the conditions can be made to optimise crystal growth.

1.5.1 Methods of protein crystallisation

There are three common methods of protein crystallisation: vapour diffusion, batch crystallisation and dialysis. The vapour diffusion method was used in this project, combined with a technique named seeding.

1.5.1.1 Vapour diffusion

Vapour diffusion is the most widely used method to produce protein crystals. A drop containing a mixture of protein and the crystallisation cocktail/precipitant is held together with the reservoir in a closed chamber. Since the precipitant is at a higher concentration in the reservoir than in the drop (adding the protein sample into the drop dilutes the precipitant), water (sometimes organic solvents) diffuses from the drop into the reservoir solution in order to reach equilibrium. As a result, the protein solution is gradually brought to supersaturation, where crystallisation occurs (Figure 1.6). The advantage of vapour diffusion method is that it is a reasonably slow process; therefore the system gradually reaches the nucleation zone rather than being quickly driven into the precipitation zone. This method is well suited for small sample volumes (0.05 – 2µl). The hanging

and sitting drop methods are frequently used techniques for crystallisation by vapour diffusion (Figure 1.6).

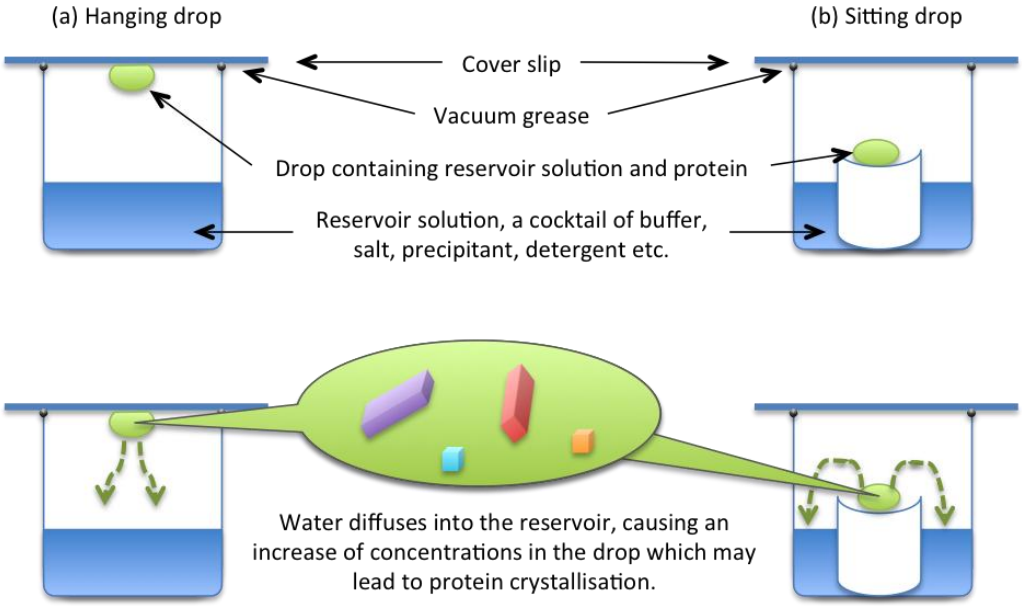


Figure 1.6. A schematic representation of the most common vapour diffusion methods for protein crystallisation. (a): Hanging drop method; (b): Sitting drop method.

Seeding

Seeding is a technique where a small crystal seed is transferred into a pre-equilibrated drop to allow free protein molecules in the solution to deposit on the surface of the seed and grow large crystals. The introduction of a crystal seed reduces random intermolecular interactions and drives such interactions to a more ordered form. Commonly, seeding can be achieved by two methods: micro-seeding and macro-seeding. In the former method, a dilution series of broken fragments of protein crystals is prepared and a portion of the diluted solution is added to a new crystallisation drop; in the latter method, a single and well-formed crystal (that is too small to directly collect data from) is used as seed. In both methods, the new crystallisation drop should be pre-equilibrated and the protein should be at a slightly lower concentration than the sample used to produce the seed (Rupp, 2010).

1.5.1.2 Batch crystallisation

Batch crystallisation is the classic technique of protein crystallisation. Relatively high concentrations of protein and crystallisation solution are mixed together so that supersaturation is reached immediately. Batch crystallisation is simple to perform, and the development of micro-batch method allows the use of small sample volumes.

1.5.1.3 Dialysis

In dialysis, the protein solution is separated from the crystallisation solution by a semi-permeable membrane, which allows water and small molecules to be freely exchanged and at the same time prevents protein molecules from circulating. Although large volumes of protein samples are required, dialysis often gives rise to large crystals.

1.6 Crystals and symmetry

A crystal is a periodical arrangement of a regular repeating unit in all three dimensions. This repeating unit is known as the unit cell, and represents the basic building block of the crystal, from which the whole crystal can be built (Figure 1.7). A unit cell is defined by its three edges designated as a, b, and c, as well as three angles α (the angle between b and c), β (the angle between a and c), and γ (the angle between a and b) (Figure 1.7).

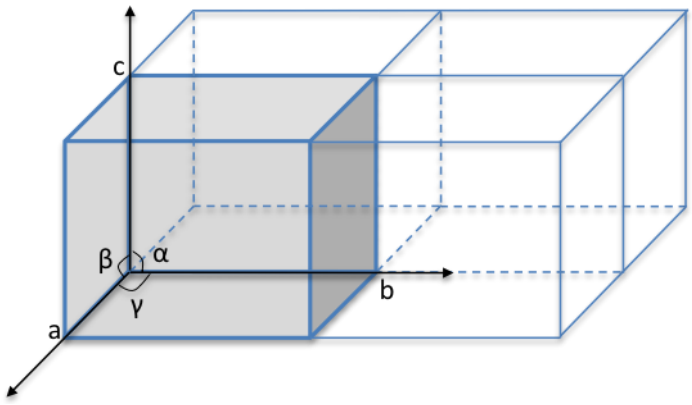


Figure 1.7. Diagrammatic representation of the dimensions and angles of a unit cell (shaded grey), as well as the arrangement of unit cells in a crystal.

Based on the type and property, a crystal can be grouped into one of the seven crystal systems (Blundell and Johnson, 1976; Blow, 2002; Rupp, 2010) (Table 1.2). These seven systems are primitive lattices (designated by letter P), containing one point at each corner of the unit cell. In addition to these seven primitive lattices, Bravais discovered another seven non-primitive lattices. These non-primitive lattices can be categorised into three types: if there is an extra lattice point in the centre of the unit cell in addition to the points at the corners, it is classified as body centred and designated by letter I. If the cell has two extra lattice points in the centre of one pair of opposite faces, or in the centre of all lattice faces, it is classified as face centred, and designated by letter C or F respectively (Table 1.2). The seven primitive lattices together with the seven non-primitive lattices are called Bravais lattices (Blundell and Johnson, 1976; Blow, 2002; Rupp, 2010).

There are four types of symmetry operators in general: rotation, translation, mirror plane, and inversion, and the combination of these symmetry operators forms 32 crystallographic point groups. The 14 Bravais lattices together with the 32 point groups give rise to 230 space groups. For asymmetric, chiral molecules such as proteins, it is necessary to maintain their handedness, which eliminates mirror plane and inversion operations, resulting 65 chiral space groups (Blundell and Johnson, 1976; Rupp, 2010).

Table 1.2. The seven crystal systems and fourteen Bravais lattices.

Crystal System	Constraints on interaxial angles and axial lengths	Bravais Lattices	Minimum internal symmetry
Triclinic	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$	P	None
Monoclinic	$a \neq b \neq c$ $\alpha = \gamma = 90°, \beta \neq 90°$	P C	2-fold rotation axis parallel to b
Orthorhombic	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$	P C I F	Three perpendicular, non-intersecting 2-fold axes
Tetragonal	$a = b \neq c$ $\alpha = \beta = \gamma = 90°$	P I	4-fold rotation axis parallel to c
Trigonal	$a = b \neq c$ $\alpha = \gamma = 90°, \beta = 120°$	P	3-fold rotation axis parallel to c
Hexagonal	$a = b \neq c$ $\alpha = \gamma = 90°, \beta = 120°$	P	6-fold rotation axis parallel to c
Cubic	$a = b = c$ $\alpha = \beta = \gamma = 90°$	P I F	Four 3-fold axes along space diagonals

1.7 Data collection

1.7.1 Cryo-protection of protein crystals

Protein diffraction data collection is now routinely carried out under cryogenic conditions. This reduces radiation damage to the crystal, allows a complete dataset to be collected from a single crystal, and in turn lowers systematic errors and improves data quality (Rodgers, 1994). In cryo-crystallography, a protein crystal is mounted onto a litho-loop, flash-frozen with and kept in liquid N_2 until data collection. To prevent the formation of crystalline ice during flash cooling, one or more species of cryo-protectant is added to the drop prior to crystal mounting. In general, there are two types of cryo-protecting agents: ① cryo-protectants such as glycerol and ethylene glycol, can penetrate into the solvent channels of crystals, preventing the formation of crystalline ice; ② cryo-protectants, for example oil, that do not travel into the crystal but act as a barrier between the crystal and the surrounding environment (Garman and Owen, 2007). Some common cryo-protectants include glycerol, ethylene glycol, polyethylene glycol (PEG), methylpentanediol (MPD), and certain lithium and carboxylate salts (Vera and Stura, 2014).

1.7.2 Diffraction data collection

Flash-cooled protein crystals are now ready to be exposed to an X-ray source for data collection, which can be completed with either in-house X-ray sources or synchrotron radiation sources. In-house X-ray sources produce X-rays by bombarding high velocity electrons onto a metal target in vacuum. Copper is commonly used in the generation of X-rays, with a wavelength of 1.54 Å. If shorter wavelengths are required, molybdenum (wavelength = 0.71 Å) can be used (Blundell and Johnson, 1976). In-house X-ray sources are convenient to use; however, the intensity of the emitted beam is very low and diffraction images have to be recorded for a very long time. Nowadays, synchrotron radiation sources are more routinely used for X-ray diffraction data collections.

In a synchrotron, electrons are produced in an electron gun, accelerated in a linear accelerator, boosted in energy, and fed into the storage ring. The electrons are forced to the circular orbit by a series of bending magnets. These magnets cause the electrons to lose energy in the form of light, which can be channelled out of the storage ring and into the experimental stations (Figure 1.8). Synchrotron sources offer a wide range of wavelengths for different experiments, and thanks to the high-energy X-ray beams, diffraction data with high signal-to-noise ratio can be collected with shorter exposure times.

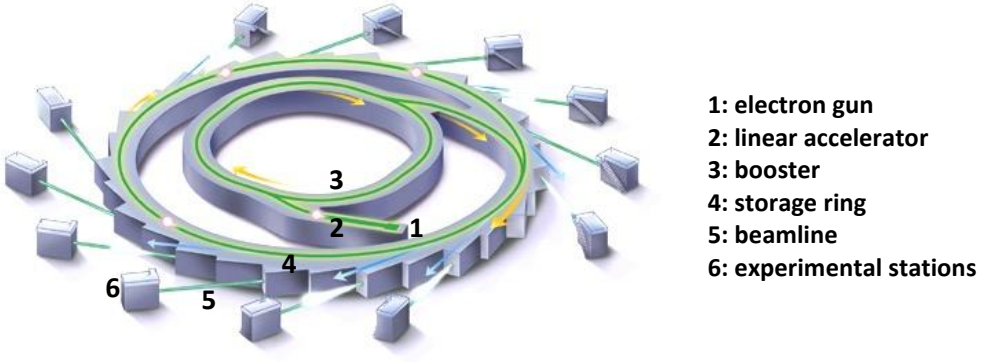


Figure 1.8. Schematic representation of a synchrotron source. Figure adapted from http://www.synchrotron.org.au/synchrotron-science/what-is-a-synchrotron.

1.8 Principles of X-ray diffraction

More than 100 years ago, William Lawrence Bragg and William Henry Bragg proposed a theory that has now become the fundamental basis of X-ray crystallography (Bragg, 1914; Bragg, 1924). This is Bragg's law, and refers to the simple equation:

$$n\lambda = 2d \sin\theta$$

The Braggs modelled the crystal as a set of parallel planes that are defined by a set of indices (*h k l*), the Miller indices. These parallel planes are separated by a distance d. For two incident X-rays with an incidence angle θ to be still in phase after scattering, the extra distance travelled by the second wave must equal one or multiples of its wavelength that is denoted as nλ, where n is an integer (Figure 1.9). Such in phase waves would constructively interfere and produce a strong diffraction maximum.

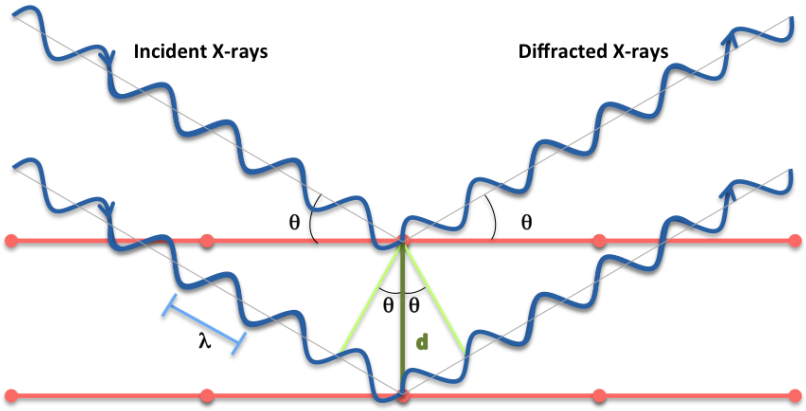


Figure 1.9. A schematic representation of Bragg's Law.

The structural information that one can obtain from X-ray diffraction data depends on the resolution of the data. Resolution is defined as the minimum interplanar spacing *d* for a particular set of X-ray diffraction intensities. The higher the resolution, the smaller the spacing, and the more detailed and accurate the structure is. Table 1.3 shows the level of detail of a structure one could expect at a given resolution.

Table 1.3. Relationship between resolution and structural information obtained.

Resolution	Structural information obtained
6.0 Å	The outline of the molecule, the quaternary structure or the domain arrangement of a macromolecule can be resolved. Sometimes the secondary structures (α-helices and β-strands) are visible.
3.0 Å	Course of the polypeptide chain is traceable; the topology of the protein can be established. The side chains of residues can be built into the model with the aid of amino acid sequence.
2.0 Å	Main chain conformations of the molecule can be unambiguously established. Conformations of side chains, some bound water molecules or solvents are recognisable.
1.5 Å	Individual atoms are almost resolved.
1.0 Å	Hydrogen atoms may become visible; alternative conformations of side chains may become visible.

The term structure factor (denoted as F_{hkl}) describes the resultant wave from all scattered waves produced by all atoms in the unit cell. For a unit cell containing j atoms, the structure factor is expressed as:

$$F_{hkl} = \sum_{j} f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

The main objective of X-ray crystallography is to determine $\rho(x, y, z)$, the electron density for all x, y, z coordinates in the unit cell. The electron density is the Fourier transformation of structure factors, and can be expressed as:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} |F_{hkl}| \, e^{i\alpha_{hkl} - 2\pi i(hx + ky + lz)}$$

where V is the volume of the unit cell, (x, y, z) are the grid position coordinates, and h, k, l are the Miller indices.

The amplitude of the structure factors $|F_{hkl}|$ are directly related to the diffracted intensities: $I \propto |F_{hkl}|^2$ therefore can be obtained directly from the experimental data. However, the phase of the structure factors α_{hkl} cannot be recorded, and is lost during data collection, without which the electron density map of the structure cannot be built. This is known as the phase problem. There are several

methods available for overcoming the phase problem, such as molecular replacement (MR) (Rossmann and Blow, 1962), isomorphous replacement (IR) (Green *et al.*, 1954), and anomalous diffraction (Hendrickson, 1991).

The raw diffraction data are processed with softwares to determine the space group and the unit cell parameters of the crystal, and to index, integrate, scale, and merge the reflections. The quality of the merged data can be assessed by the linear merging R-value R_{merge}:

$$R_{merge} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - < I_{hkl} > |}{\sum_{hkl} \sum_j I_{hkl,j}}$$

where $< I_{hkl} >$ is the average intensity of each reflection. A good set of diffraction data should have an R_{merge} value of ~4–5%; however, one needs to bear in mind that it is not the most useful indicator, as it does not take the redundancy of the diffraction data into account (lower redundancy will give lower R_{merge}) (Wlodawer *et al.*, 2008; Rupp, 2010). The signal-to-noise ratio (SNR) $< I/\sigma(I) >$ provides an estimate for the usefulness of the data: usually the diffraction limit is set to a resolution where the SNR drops below 2.0. The redundancy and completeness are also useful indicators of the quality of the diffraction data. In principle, both high redundancy and high completeness are desirable, as the former improves the quality of the merged data, whereas the latter reduces the noise level and enhances the interpretability of calculated maps (Wlodawer *et al.*, 2008; Rupp, 2010).

Once the phases are obtained, an initial molecular model can be built into the calculated electron density map with the help of the amino acid sequence as well as any known structural characteristics of the target protein. The initial model is far from perfect, and it needs to be refined to improve the agreement between the experimental diffraction data and the calculated phases. The quality of the model is measured by calculating the R-factor, a measure of the level of agreement between diffraction data and the model:

$$R = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|}$$

where $|F_{obs}|$ is the amplitudes of the observed structure factor, and $|F_{calc}|$ is the amplitudes of the calculated structure factor of the model.

The higher the agreement between the observed and calculated amplitudes, the closer the R factor is towards zero. However, the R-factor can be made small by increasing the number of model parameters, which would over-fit the diffraction data and result in a model that is physically incorrect (Rupp, 2010). To overcome this bias, an additional factor called R_{free} is introduced (Brünger, 1992). R_{free} is calculated the same way as the R-factor using 5-10% of the total reflections that are excluded from refinement. Any over-fitting of the model is reflected in an increase of the R_{free} value.

Chapter 2. Introduction to the ribonuclease A (RNase A) superfamily

2.1 Introduction

Early molecular biology studies revealed that biological information is stored in deoxyribonucleic acid (DNA) and fulfilled by proteins, with ribonucleic acid (RNA) serving as the link between these two types of molecules (Crick, 1970).

$$DNA \leftrightarrow RNA \rightarrow protein$$

The essential flow of information through RNA is controlled by its synthesis or degradation. This involves two classes of enzymes. RNA polymerases catalyse the synthesis of RNA, whilst RNA degradation is achieved by the activities of RNA depolymerases, most commonly known as ribonucleases (RNases) (Raines, 1998).

RNases catalyse the cleavage of the phosphodiester (P-$O^{5'}$) bond of RNA, and depending on the binding position on the RNA molecule, RNases can be subcategorised into exo- and endo-ribonucleases. Exoribonucleases, for instance RNase D, carry out RNA degradation by removing nucleotides at either the 5' or the 3' end of the RNA molecule. In contrast, endoribonucleases, such as RNase A and RNase P, catalyse the cleavage of phosphodiester bonds within a polynucleotide RNA chain (Gegenheimer and Apirion, 1981).

RNases are involved in the degradation of many RNA species, including the three most abundant RNA species in all organisms: ribosomal RNA (rRNA), messenger RNA (mRNA), and transfer RNA (tRNA). RNases have been found in all kingdoms of life, from microorganisms to mammals, demonstrating the importance of RNA regulation in cellular processes in all organisms.

Extensive studies on RNases have provided researchers with extensive knowledge regarding their evolution, catalytic mechanisms, structures, and cellular functions. Interestingly, recent studies have also revealed special

biological properties of some RNases, such as cytotoxicity, the ability to induce and promote angiogenesis and neurogenesis, and roles in host defence. These make them potential targets for novel therapeutic agents. In this regard, our knowledge of bovine pancreatic ribonuclease (RNase A), the prototype of the RNase A superfamily, could be used to shed light on the physical, chemical and biological properties of these RNases.

2.2 Bovine pancreatic ribonuclease (RNase A), the prototype of ribonucleases

The study of the RNase A superfamily started with the discovery (Jones, 1920) and isolation (Dubos, 1937) of bovine pancreatic ribonuclease (EC = 3.1.27.5) in 1920 and 1937 respectively. The extensive studies on the folding, structure, and chemistry of RNase A that have followed have lead to the awarding of the Nobel Prize in Chemistry to four researchers in 1972 and 1984, thereby demonstrating the importance of this enzyme (Anfinsen, 1973; Moore and Stein, 1973; Merrifield, 1986).

The enzyme was given the name RNase A to refer to its predominant state in the pancreas of *Bos taurus* (Raines, 1998). RNase A is unmodified, whereas RNase B, C, and D, which are also isolated from bovine pancreas, are glycosylated at different sites (Plummer and Hirs, 1964; Plummer, 1968; Baynes and Wold, 1976; Liang *et al.*, 1980; Rudd *et al.*, 1992; Fu *et al.*, 1994).

RNase A is a secreted protein with a signal peptide of 26 amino acids, which is cleaved after secretion by the exocrine cells of the bovine pancreas. The mature protein has 124 amino acids with a molecular weight of 13,686 Da (Raines, 1998).

2.2.1 Reaction catalysed by RNase A

The depolymerisation of RNA catalysed by RNase A consists of two reactions: the transphosphorylation reaction and the hydrolysis reaction. During the first step (transphosphorylation), the phosphodiester bond is broken at the 5' position of one nucleotide, and the phosphate group is transferred to the 2' position of the adjacent nucleotide forming a 2',3'-cyclic phosphodiester intermediate product (Figure 2.1a). This intermediate is hydrolysed to a 3' nucleotide in the second step (hydrolysis) (Figure 2.1b) (Richards and Wyckoff, 1971; Eftink and Biltonen, 1987).

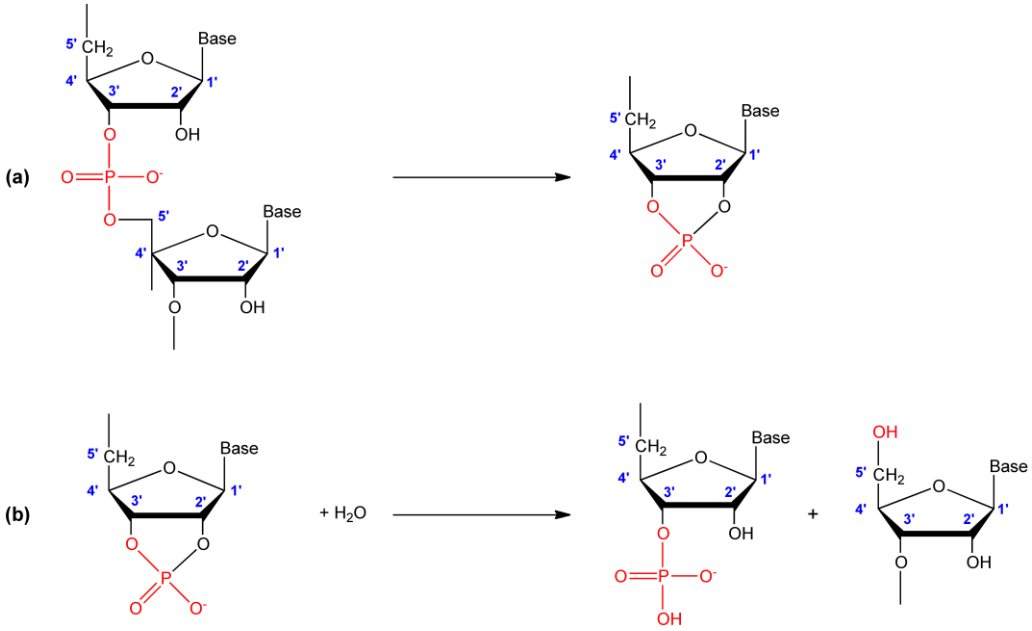


Figure 2.1. Hydrolysis reaction of RNA. (a) The transphosphorylation reaction; (b) The hydrolysis reaction.

2.2.2 Structure of RNase A

The successful crystallisation of RNase A protein was first reported in the late 1930s (Kunitz, 1939; Kunitz, 1940), and showed diffraction up to 2 Å (Fankuchen, 1941). The three-dimensional structure of RNase A was only the fifth protein structure to be solved using X-ray crystallography (Kartha *et al.*, 1967; Richards and Wyckoff, 1973), following myoglobin (Kendrew, 1963), haemoglobin (Perutz, 1963), lysozyme (Blake *et al.*, 1965), and carboxypeptidase A (Lipscomb *et al.*, 1966).

2.2.2.1 Structural features of RNase A

The overall shape of RNase A resembles that of a kidney, with a predominant secondary structural feature of a four-stranded antiparallel β-sheet. The structure is stabilised by four disulphide bonds (formed between Cys-26 and Cys-84, Cys-40 and Cys-95, Cys-58 and Cys-110, and Cys-65 and Cys-72), two at each pole of the molecule. The active site residues are located in the cleft (Figure 2.2).

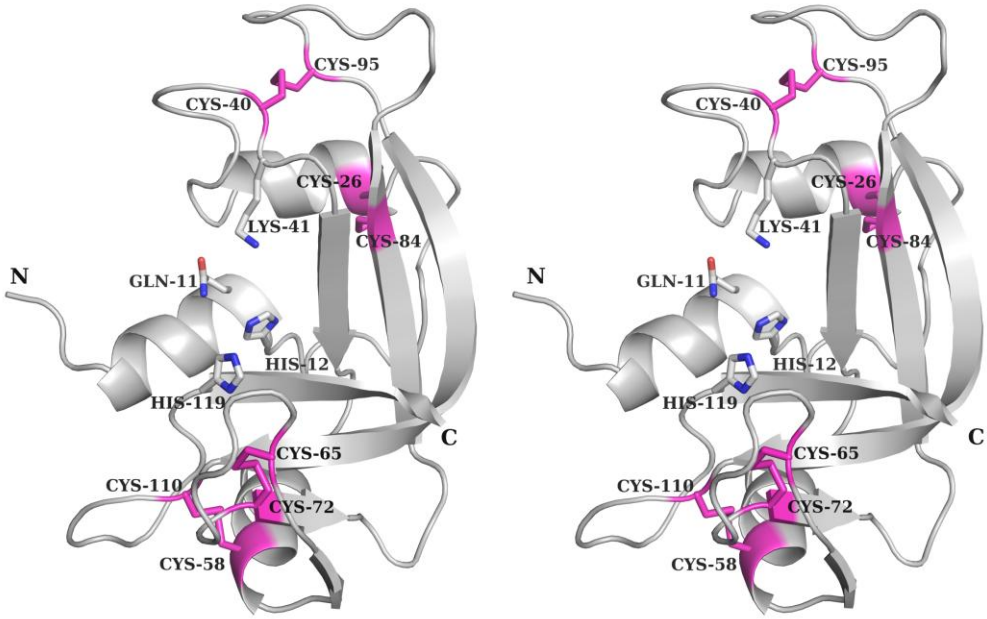


Figure 2.2. Stereo view of the three-dimensional structure of bovine pancreatic ribonuclease A (RNase A) (PDB: 1AFU; Leonidas *et al.*, 1997). Four pairs of cysteine residues involved in intra-molecular disulphide bonds are shown in pink; active site residues (Gln-11, His-12, Lys-41, and His-119) are shown as sticks.

2.2.2.2 Catalytic mechanism of RNase A

Researchers have proposed a few catalytic mechanisms for RNase A (Witzel, 1963; Hammes, 1968; Wang, 1968), one of which is depicted in Figure 2.3. This mechanism, which is consistent with all known data from research performed on this enzyme, postulates that two histidine residues participate in the catalysis of RNA hydrolysis (Findlay *et al.*, 1961). During the

transphosphorylation step, the side chain of His-12 acts as a base, abstracting the proton on the 2'-O of one nucleotide. The abstraction leads to further electron transfer along the nucleotide molecule, resulting in the attack of 2'-O on the phosphorus atom and the formation of the 2',3'-cyclic phosphodiester intermediate product (Usher *et al.*, 1970; Usher *et al.*, 1972; Sowa *et al.*, 1997). The side chain of His-119 acts as an acid and facilitates the displacement of 5'-O (Figure 2.3a). The second step, the slow hydrolysis of the 2',3'-cyclic phosphodiester occurs in a separate process that resembles the reverse of transphosphorylation (Cuchillo *et al.*, 1993; Thompson *et al.*, 1994). Here the roles of the histidines are reversed: the side chain of His-119 donates an electron to a water molecule, which subsequently passes the electron to the phosphorus atom and finally to the nitrogen atom of the side chain of His-12, producing the final reaction product (Figure 2.3b). DNA lacks the 2'-OH group, and therefore cannot be hydrolysed by RNases.

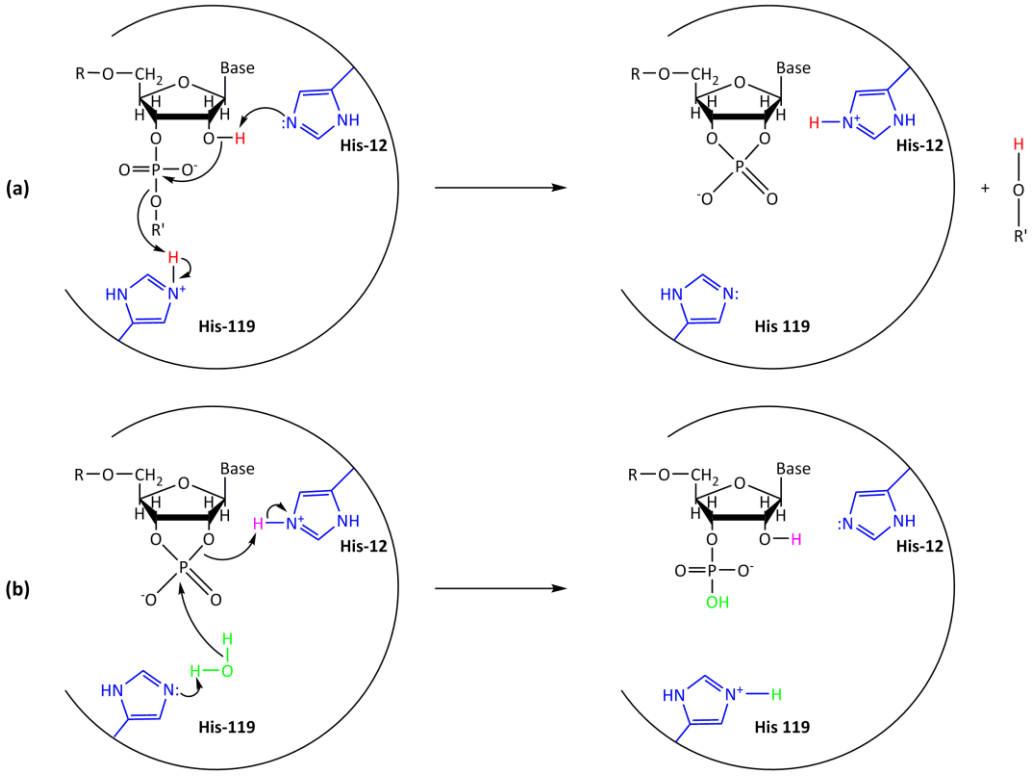


Figure 2.3. Catalytic mechanism of RNase A. (a) The transphosphorylation step where an intermediate product, 2',3'- cyclic phosphodiester, is produced; (b) The hydrolysis reaction of the intermediate product.

Both the transphosphorylation and hydrolysis of RNA catalysed by RNase A occur via transition states having a pentavalent phosphorus atom (Raines, 1998). Uridine 2',3'-cyclic vanadate (UVC; Figure 2.4a), an analogue of the penta-coordinate phosphorane bearing some resemblance to such a transition state (Lindquist *et al.*, 1973), has been used to study the catalytic mechanism of RNase A (Borah *et al.*, 1985; Ladner *et al.*, 1997) (Figure 2.4b). Potential hydrogen bonds are identified between UVC and Gln-11, His-12, Lys-41, Thr-45, His-119, and Phe-120 of RNase A (Figure 2.4c).

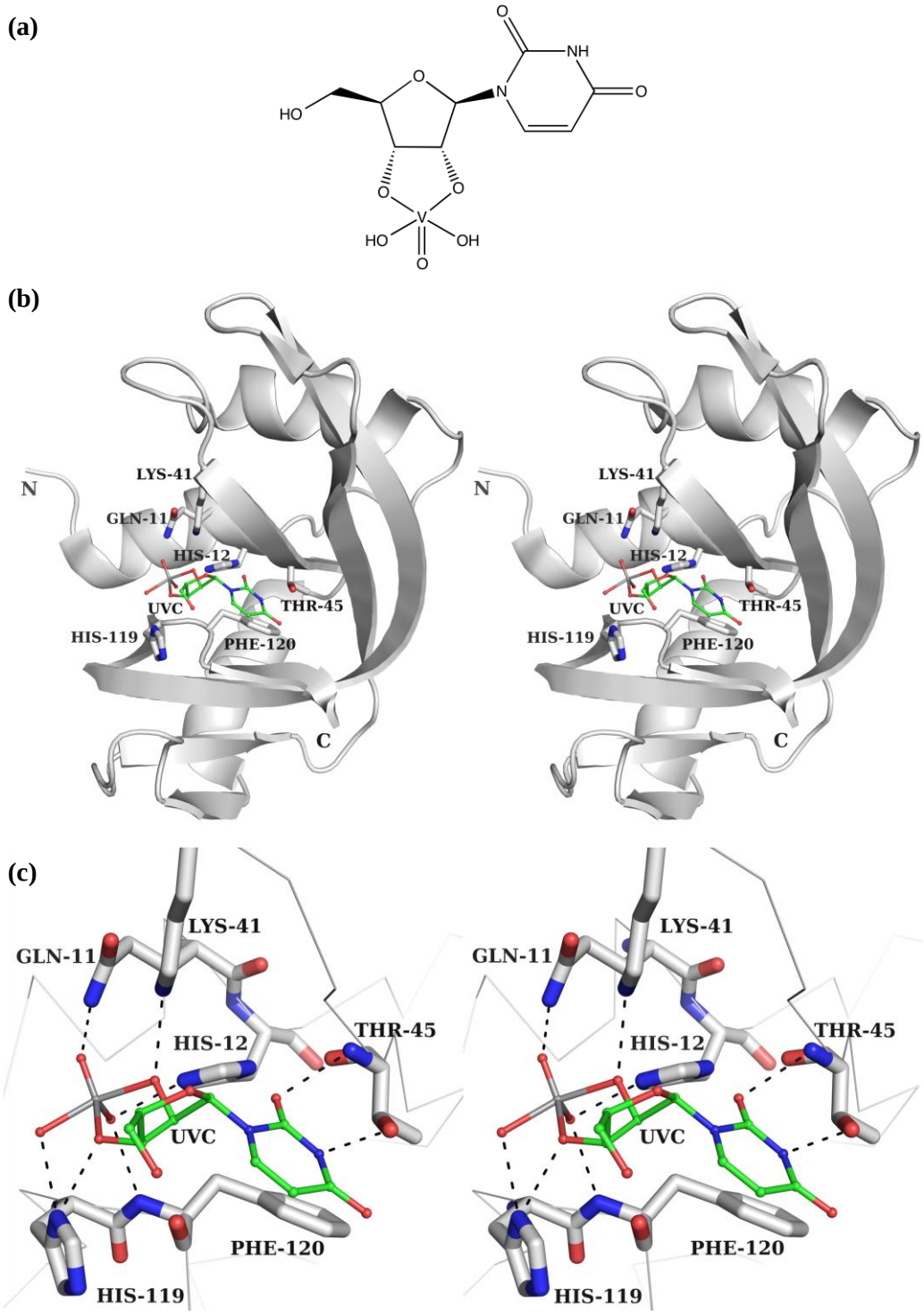


Figure 2.4. Interactions between RNase A and uridine 2',3'-cyclic vanadate (UVC). (a) Chemical structure of UVC; (b) Stereo view of the structure of RNase A-UVC complex (PDB: 1RUV; Ladner *et al.*, 1997); (c) Stereo view of the interactions between UVC and RNase A. UVC and the side chains of the residues involved in the recognition of UVC are shown as sticks; the N- and C- termini of RNase A are labelled as N and C respectively. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms of RNase A and UVC are coloured in grey and green respectively; potential hydrogen bonds are shown as dashed lines.

2.2.2.3 Catalytic site of RNase A

The crystal structures of RNase A complexed with UVC have provided invaluable insight into the catalytic mechanism of RNase A, and revealed that the side chains of Gln-11, His-12, Lys-41, His-119 are proximal to the vanadyl group and therefore of catalytic importance (Wlodawer *et al.*, 1983; Ladner *et al.*, 1997) (Figure 2.4b).

A close view of the active site is shown in Figure 2.5. His-12 and His-119 were identified as important catalytic residues from early work on RNase A, and their direct role in RNA degradation has been confirmed (Figure 2.3). As stated previously, His-12 acts as a base in the transphosphorylation reaction that abstracts a proton from the 2'-O of the ribose, thus favouring the attack of 2'-O on the phosphorus atom to give the pentacoordinate transition state (Cuchillo *et al.*, 1997). His-119, on the other hand, acts as an acid in the transphosphorylation, protonating the 5'-OR leaving group and facilitating an in-line addition of 2'-O (Deakyne and Allen, 1979). Several crystallographic studies have shown that His-119 occupies predominantly two conformations (Richards and Wykoff, 1971; de Mel *et al.*, 1994), and modelling studies have shown that His-119 is positioned equally well to participate in the RNA degradation reactions regardless of its conformation (de Mel *et al.*, 1992).

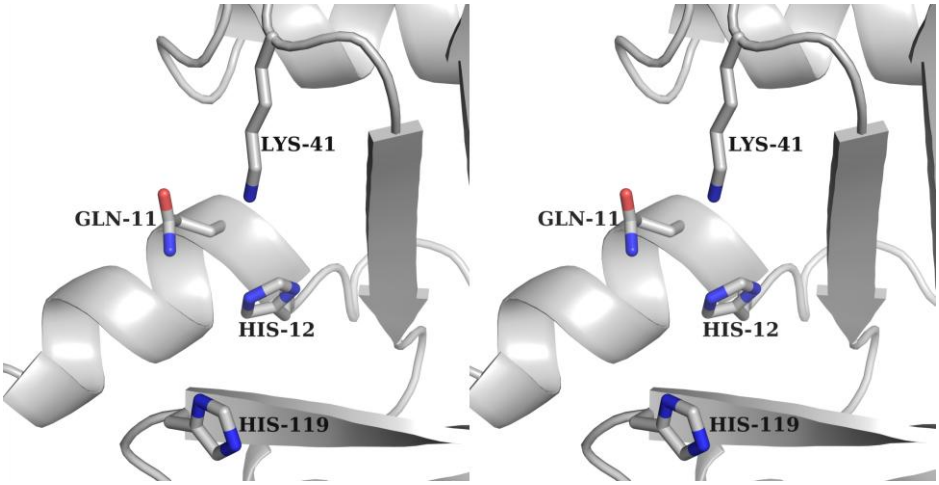


Figure 2.5. Close view of the active site of RNase A. Amino acid residues of catalytic importance are shown as sticks. Carbon, nitrogen and oxygen atoms are coloured in grey, blue and red respectively.

Early chemical modification of Lys-41 suggested that it may contribute to the catalytic activity of RNase A (Hirs, 1962; Murdock *et al.*, 1966), and this was later confirmed by site-directed mutagenesis studies on Lys-41. The K41A variant showed 1.9% activity with respect to the wild-type RNase A (Trautwein *et al.*, 1991).

The role played by Lys-41 in the catalytic activity of RNase is to stabilise the excess negative charge accumulated on the non-bridging phosphoryl oxygens in the transition state. A few hypotheses were made on the mechanism of this stabilisation (Roberts *et al.*, 1969; Flogel *et al.*, 1975; Trautwein *et al.*, 1991), and a model where Lys-41 donates a single hydrogen bond to the transition state is supported by available data (Figure 2.6) (Raines, 1998).

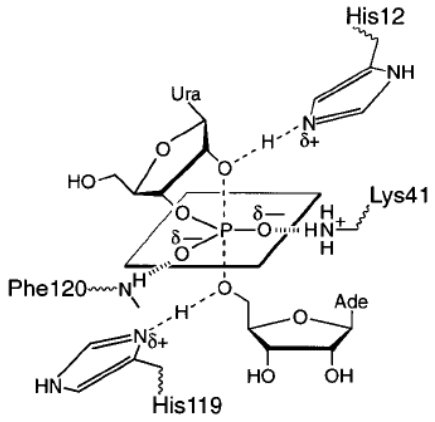


Figure 2.6. Putative structure of the transition state of the transphosphorylation of UpA by RNase A. Figure adapted from "Ribonuclease A" (Raines, 1998).

Analysis of RNase A structures revealed that the side chain of Gln-11 can form a hydrogen bond to a substrate/substrate analogue (Figure 2.4c), phosphate ion or sulphate ion in the active site (Gilliland, 1997; Raines, 1998). Further mutagenesis and enzyme kinetic studies on Gln-11 revealed that a hydrogen bond formed between the side chain of Gln-11 and a phosphoryl oxygen prevents the substrate from binding in a non-productive mode, and enhances the catalysis of RNA hydrolysis (delCardayré *et al.*, 1995; Raines, 1998).

2.2.2.4 Subsites of RNase A and substrate specificity

Structural studies of RNase A not only revealed the primary active site consisting of amino acid residues Gln-11, His-12, Lys-41, and His-119, but also identified so called subsites that are involved in the recognition of other components of an RNA molecule in the catalytic cleft (Figure 2.7a). So far, three subsites denoted P_0, P_1, and P_2 have been shown to be involved in the binding of phosphate groups of RNA, whereas two further subsites, B_1 and B_2, interact with the bases of the RNA (Parés *et al.*, 1980; Parés *et al.*, 1991; Noguès *et al.*, 1995).

As shown in Figure 2.7b, the P_0 subsite consists of only one amino acid residue: Lys-66. The P_1 subsite is the primary active site where residues Gln-11, His-12, Lys-41, and His-119 catalyse the degradation of RNA, whilst residues Lys-7 and Arg-10 contribute to the binding of a phosphate group to the P_2 subsite.

The B_1 subsite comprising residues Val-43, Asn-44, Thr-45, Phe-120, and Ser-123 is a pyrimidine specific binding site, and it has been demonstrated that it has a 30-fold kinetic preference for cytosine over uracil (McPherson *et al.*, 1986; Aguilar *et al.*, 1992). In contrast, the B_2 subsite with residues Asn-67, Gln-69, Asn-71, and Glu-111 prefers the binding of purines, specifically an adenine base (Katoh *et al.*, 1986). The contributions of some amino acid residues to substrate binding and/or turnover have been verified by site-directed mutagenesis studies (Wlodawer *et al.*, 1983; Trautwein *et al.*, 1991; Tarragona-Fiol *et al.*, 1993; Thompson and Raines, 1994; delCardayré and Raines, 1994; Boix *et al.*, 1994; delCardayré and Raines, 1995; delCardayré *et al.*, 1995; Messmore *et al.*, 1995; Thompson *et al.*, 1995; Panov *et al.*, 1996; Fisher *et al.*, 1998; Schultz *et al.*, 1998).

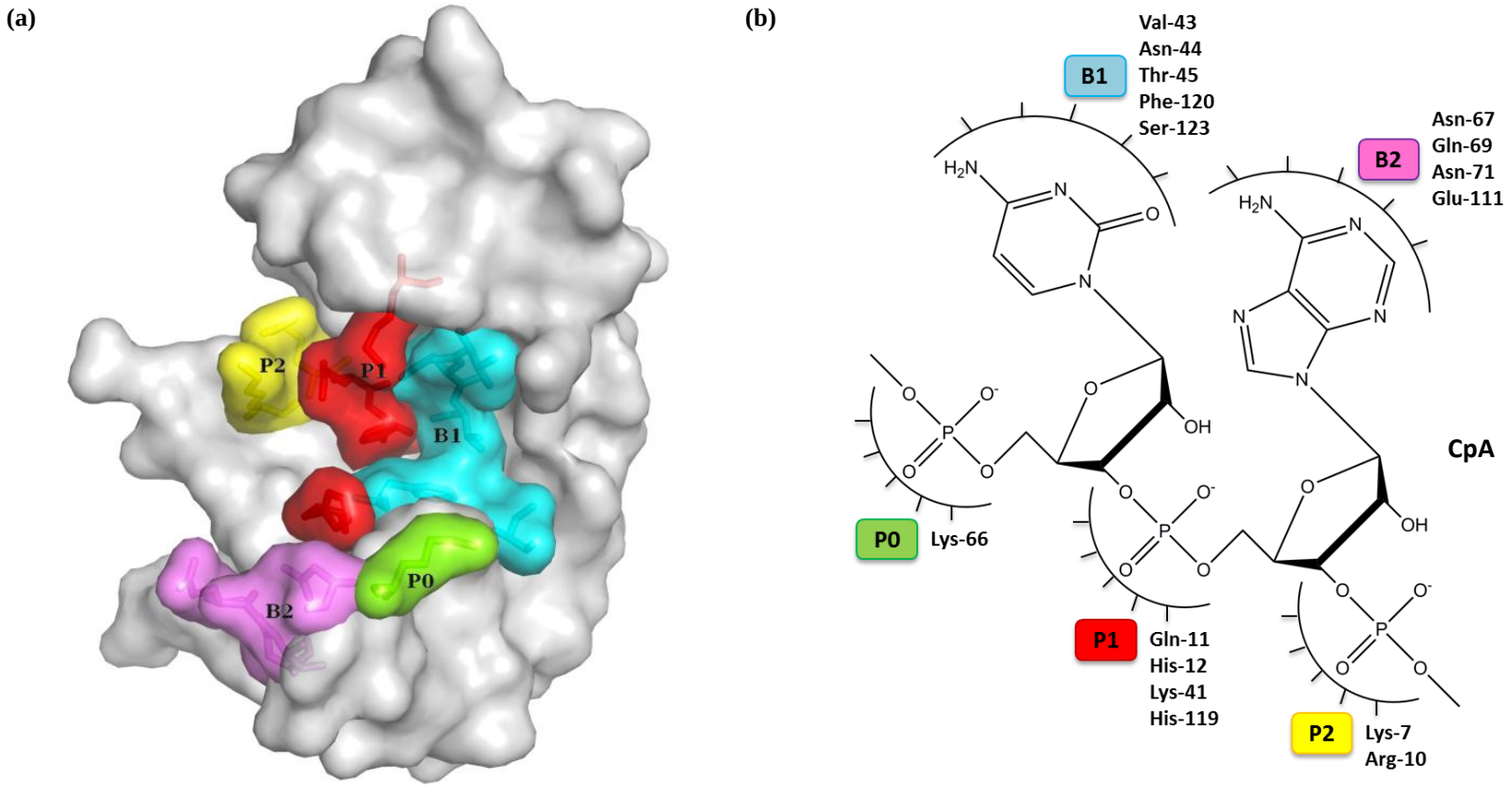


Figure 2.7. Subsites of RNase A involved in the binding of RNA. (a) Surface diagram of the three-dimensional structure of bovine RNase A showing the subsites involved in the recognition and degradation of RNA. B and P indicate the base and phosphate binding subsites respectively; (b) Schematic representation of the recognition of an RNA fragment (CpA) by RNase A. Residues in the different subsites are labelled. The phosphate group of the phosphodiester bond to be hydrolysed by the enzyme binds to P_1; B_1 is specific for pyrimidines, B_2 shows a preference for purines.

2.3 The RNase A superfamily

2.3.1 Human RNase A homologues

The analysis of human chromosome 14q11.2 revealed the complete identification of human RNase genes. A total of thirteen RNase genes were found, eight of which encode for the so-called canonical RNases (RNase 1-8). The remaining five, RNase 9-13, were named noncanonical RNases due to the lack of one or more catalytic residues (Cho *et al.*, 2005). The major differences between the canonical and noncanonical RNases lie in the C-terminal regions (Castella *et al.*, 2004; D'Alessio, 2011). As for their function, a couple of the noncanonical RNases (RNase 9 and 10) are expressed in mouse and porcine epididymis, suggesting they may play a role in the reproduction process of these species (Castella *et al.*, 2004; Penttinen *et al.*, 2003).

The structures of RNase 1-5 and RNase 7 have been solved, all of which resemble that of RNase A, the typical kidney-shaped molecule (Figure 2.8). Multiple sequence alignments of mature human RNase 1-8 with bovine RNase A revealed highly conserved active site residues and cysteine residues involved in the formation of disulphide bonds (Figure 2.9).

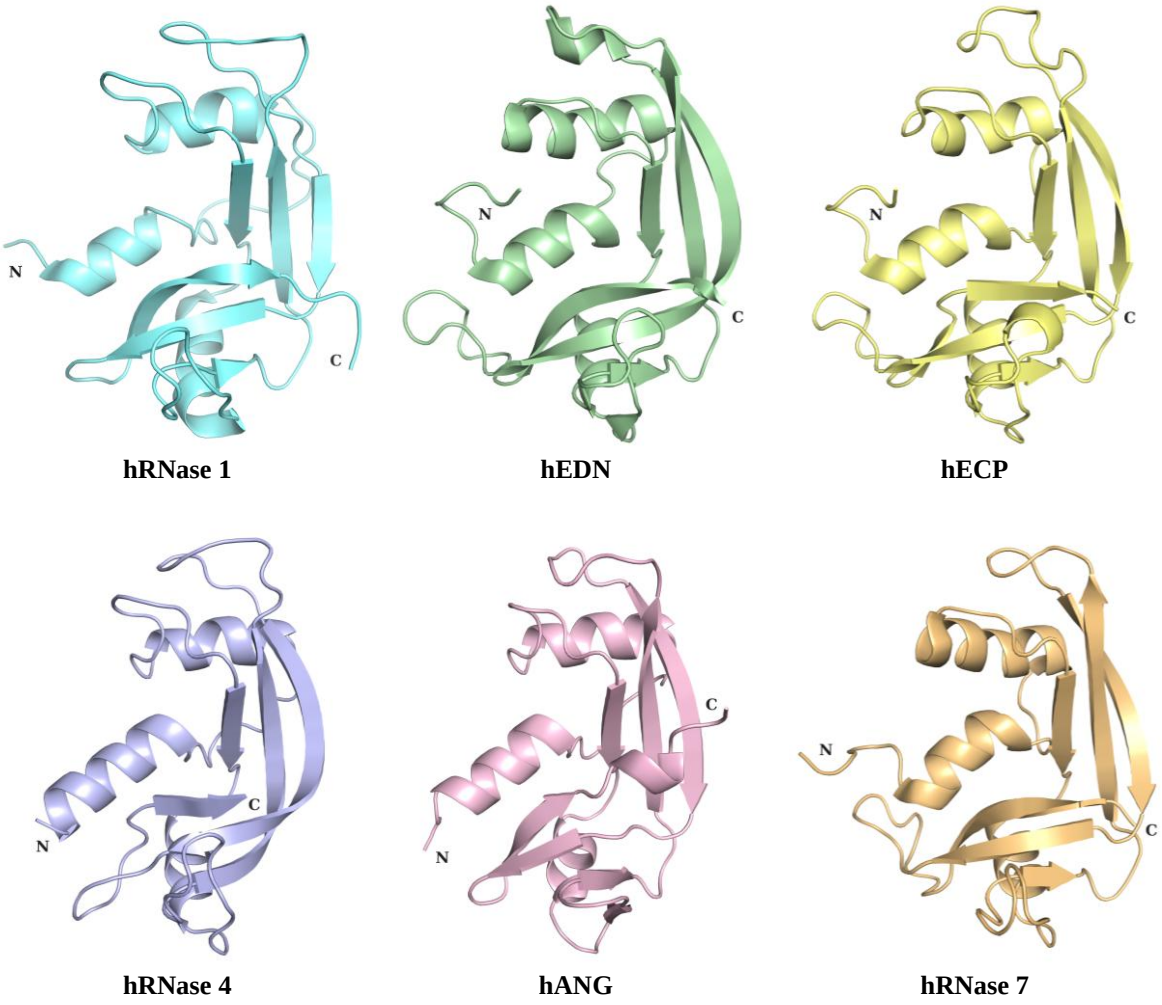


Figure 2.8. Structural similarity of human RNase A homologues – RNase 1 (hRNase 1; PDB: 2K11; Kövér *et al.*, 2008), eosinophil derived neurotoxin (hEDN; PDB: 1GQV; Swaminathan *et al.*, 2002), eosinophil cationic protein (hECP; PDB: 1QMT; Boix *et al.*, 1999), RNase 4 (hRNase 4; PDB: 1RNF; Terzyan *et al.*, 1999), angiogenin (hANG; PDB: 1B1I; Leonidas *et al.*, 1999a), and RNase 7 (hRNase 7; PDB: 2HKY; Huang *et al.*, 2007) are shown.

Figure 2.9. Comparison of the primary structures of human RNases 1-8. The sequence of bovine RNase A is shown for comparison. Conserved catalytic residues are coloured in red; cysteines involved in the formation of disulphide bonds are coloured in yellow. All sequences were obtained from UniProtKB database. The multiple sequence alignment was performed using *Clustal Omega* EMBL-EBI online server (Sievers *et al.*, 2011; McWilliam *et al.*, 2013).

2.3.1.1 Human pancreatic RNase 1 (RNase 1)

The expression of human RNase 1 has been identified in not only pancreas but also other tissues and organs, such as endothelial cells, dendritic cells, brain, liver and kidneys, to name but a few (Yasuda *et al.*, 1993; Lapteva *et al.*, 2001; Landré *et al.*, 2002). Such wide existence over a variety of tissues has led to extensive research and subsequent discoveries of some of the novel roles of RNase 1. These include the ability to degrade HIV-1 RNA thus preventing foetus/neonate vertical transmission of HIV-1 (Bedoya *et al.*, 2006), and possible participation in inflammatory and immune responses (Yang *et al.*, 2004).

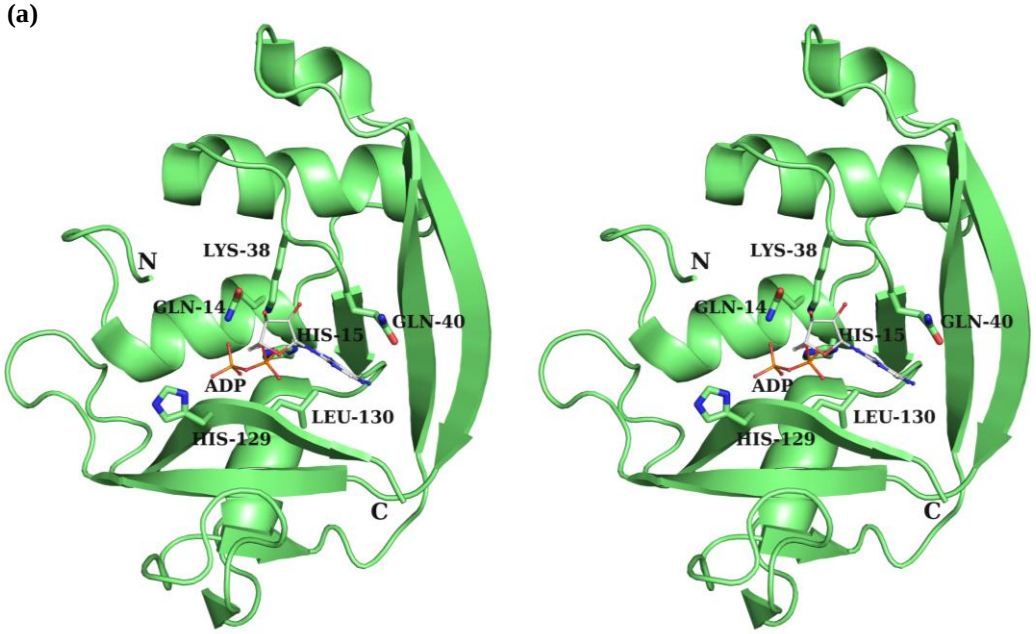
2.3.1.2 Eosinophil derived neurotoxin (EDN; RNase 2) and Eosinophil cationic protein (ECP; RNase 3)

Human RNase 2 and 3 are more commonly known as eosinophil derived neurotoxin (EDN) and eosinophil cationic protein (ECP) respectively (Figure 2.8, Figure 2.10). EDN and ECP are two of the major secretory proteins of human eosinophil granules that were first isolated prior to their identification as members of the RNase A superfamily (Ackerman *et al.*, 1983; Gleich *et al.*, 1986).

Despite the first discovery in eosinophil granules and the initial characterisation as a neurotoxin (Durack *et al.*, 1981; Fredens *et al.*, 1982; Gleich *et al.*, 1986), EDN is in fact expressed in spleen, liver, kidney, placenta, and body fluids (Beintema *et al.*, 1988; Sorrentino *et al.*, 1988; Mizuta *et al.*, 1990; Yasuda *et al.*, 1990; Shapiro and Vallee, 1991; Sorrentino *et al.*, 1992). EDN displays a specific ribonucleolytic dependent antiviral activity against respiratory syncytial virus (RSV) and HIV (Domachowske, Bonville, *et al.*, 1998; Domachowske, Dyer, *et al.*, 1998; Rosenberg and Domachowske, 2001; Bedoya *et al.*, 2006). Moreover, EDN is categorised as an alarmin, for its ability to activate and promote the chemotaxis of human dendritic cells in response to pathogen stimulation (Yang *et al.*, 2003; Yang *et al.*, 2004; Oppenheim and Yang, 2005; Acharya and Ackerman, 2014).

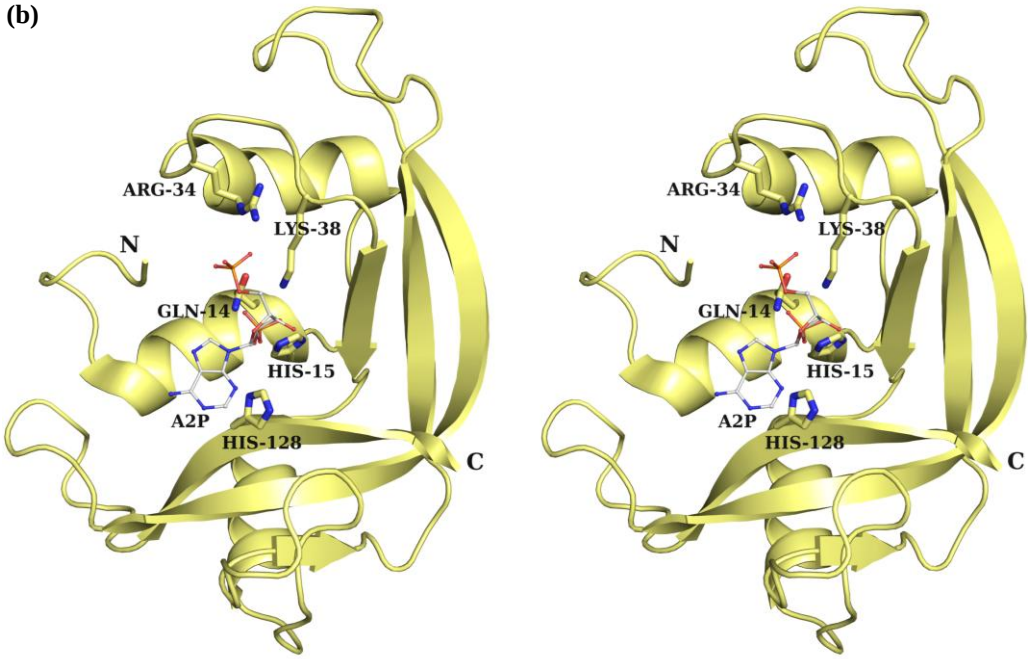
Unlike EDN, ECP is solely expressed in eosinophils activated by inflammation. ECP is recognised to have bactericidal, antiviral, and antiparasitic activities, which are independent on its ribonucleolytic activity (Yazdanbakhsh *et al.*, 1987; Lehrer *et al.*, 1989; Hamann *et al.*, 1990; Domachowske, Dyer, *et al.*, 1998). Both the cytotoxicity and bactericidal activity of ECP are based on its ability to disrupt membranes (Carreras *et al.*, 2003; Carreras *et al.*, 2005; Boix and Nogués, 2007; Navarro *et al.*, 2008; Acharya and Ackerman, 2014).

(a)



hEDN ·ADP

(b)



hECP ·A2P

Figure 2.10. Ligand recognitions of hEDN and hECP in stereo view. (a) Interactions between hEDN and adenosine 2',3'-diphosphate (ADP) (PDB: 1HI5; Leonidas *et al.*, 2001); (b) interactions between hECP and adenosine 2',5'-diphosphate (A2P) (PDB: 1H1H; Boix *et al.*, 1999a). ADP and A2P ligands as well as the side chains of residues interacting with the ligands are shown as sticks; the N- and C- termini of hEDN and hECP are labelled as N and C respectively.

EDN and ECP show unusual features relating to their catalytic properties. Neither EDN nor ECP hydrolyse the 2',3'-cyclic phosphodiesters produced in the first transphosphorylation reaction. Furthermore, despite the fact that EDN and ECP have the same base preference as the prototypic RNase A, their catalytic activities of CpA are significantly weaker than that of the RNase A (Boix *et al.*, 1999b; Holloway *et al.*, 2004; Shapiro and Vallee, 1991).

2.3.1.3 Ribonuclease 4 (RNase 4)

Initial characterisation of the members of the RNase 4 family revealed high degree of conservation, indicated by the high inter-species similarity between porcine, bovine, human, and rat enzymes. Such a level of conservation suggests a high evolutionary pressure to maintain the structure, which may imply a specialised function. RNase 4 is the shortest human canonical ribonuclease, with only 119 amino acid residues in the mature protein whereas its homologues have between 123 and 134 residues, yet its ribonucleolytic activity is close to that of RNase A. Unlike the other members in the RNase A superfamily, RNase 4 prefers uracil to cytosine as the base at the 3' side of the cleaved phosphodiester bond (Hofsteenge *et al.*, 1998b; Seno *et al.*, 1995; Shapiro *et al.*, 1986; Terzyan *et al.*, 1999; Vicentini *et al.*, 1996, 1994; Zhou and Strydom, 1993) (Figure 2.11).

Surprisingly, mutagenesis studies on its orthologue – porcine RNase 4 (pRNase 4) suggested that such strong uracil preference could be reversed (to a cytosine preference) by a single amino acid substitution mutation of Asp-80 to alanine (Hofsteenge *et al.*, 1998a). A recent discovery showed that human RNase 4 (hRNase 4) is not only able to induce angiogenesis and the formation of neurofilaments – the intermediate filaments found in neurons that are essential for the development, maintenance and function of axons (Yuan *et al.*, 2012), but also protects against hypothermia-induced neuron degeneration (Li *et al.*, 2013). More detailed discussions on RNase 4 will follow in Chapter 3.

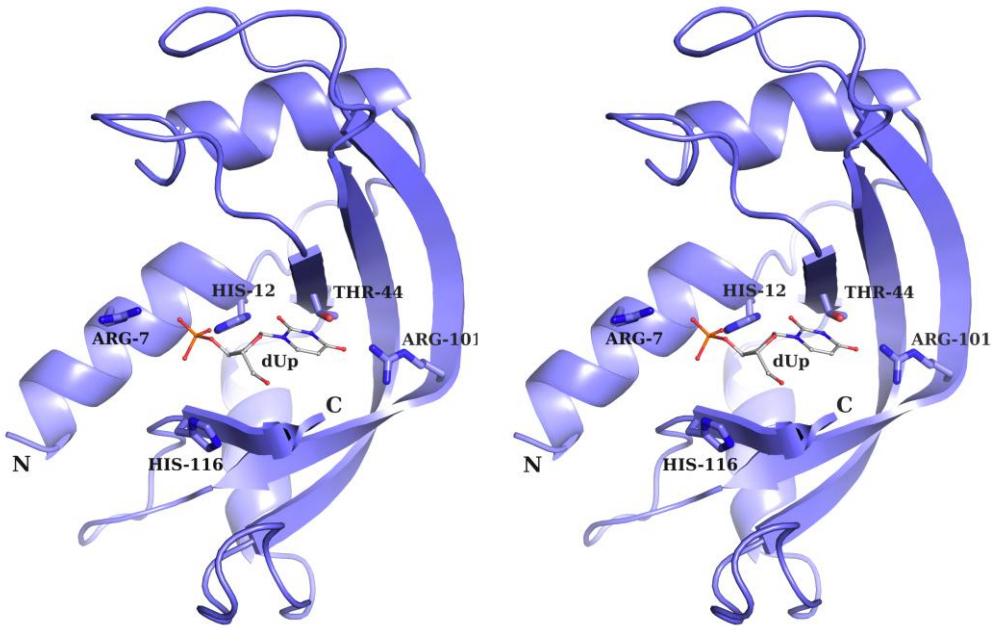


Figure 2.11. Stereo view of the structure of hRNase 4 in complex with 2'-deoxyuridine 3'-monophosphate (dUp) (PDB: 2RNF; Terzyan *et al.*, 1999). dUp as well as the side chains of residues interacting with the ligand are shown as sticks; the N- and C- termini of hRNase 4 are labelled as N and C respectively.

2.3.1.4 Angiogenin (ANG; RNase 5)

Angiogenin (ANG) is the most unusual member of the RNase A superfamily. ANG, originally isolated from tumour cell conditioned culture medium, possessed the ability to induce angiogenesis (the formation of new blood vessels) *in vivo* (Strydom *et al.*, 1985; Fett *et al.*, 1985). It exhibits a weak ribonucleolytic activity that is essential for angiogenesis (Shapiro *et al.*, 1986). Structural studies on human (Acharya *et al.*, 1994), bovine (Acharya *et al.*, 1995), and murine (Holloway *et al.*, 2005; Crabtree *et al.*, 2007) ANG revealed an overall structure similar to that of RNase A, and highlighted the less conserved B_1 and B_2 subsites, providing an explanation for the low level of ribonucleolytic activity (Figure 2.12).

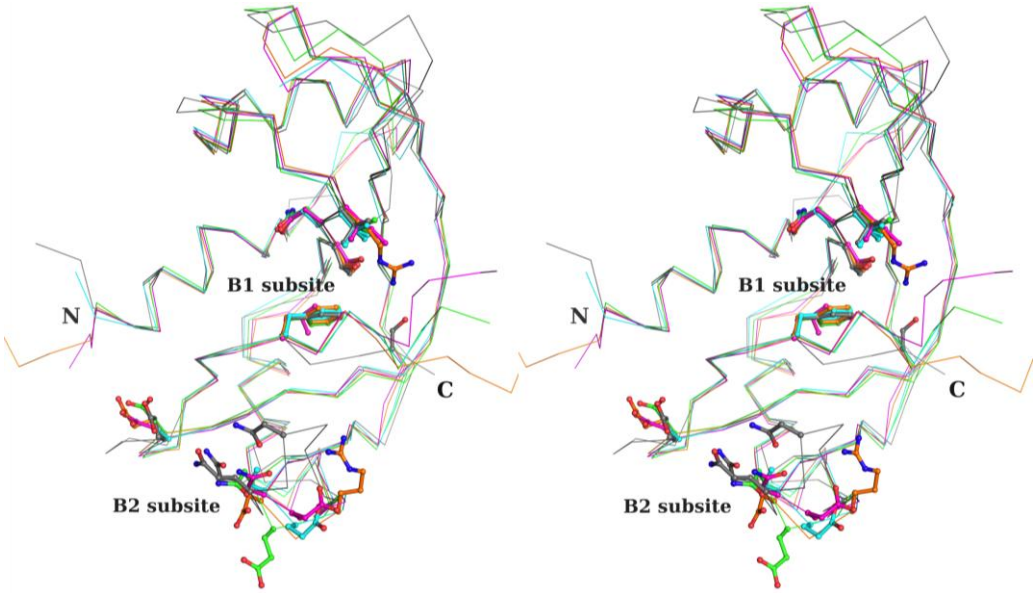


Figure 2.12. The less conserved B_1 and B_2 subsites of human ANG (magenta) (hANG; PDB: 1B1I; Leonidas *et al.*, 1999a), bovine ANG (orange) (bANG; PDB: 1AGI; Acharya *et al.*, 1995), murine ANG-1 (cyan) (mANG-1; PDB: 2BWK; Holloway *et al.*, 2005), and murine ANG-4 (green) (mANG-4; PDB: 2J4T; Crabtree *et al.*, 2007) compared with those of RNase A (grey).

The angiogenic activity of ANG is known to be associated with the pathology of cancers: the increased expression of ANG has been correlated with the formation of tumours in more than 20 clinical studies (Kao *et al.*, 2002). Studies have shown that ANG antagonists could inhibit the establishment, progression, and metastasis of tumour xenografts in athymic mice (Kao *et al.*, 2002; Olson *et al.*, 1995, 1998, 2002; Piccoli *et al.*, 1998).

The biological role of ANG is not only found in angiogenesis, but also extends to the nervous system. ANG shows neuroprotective activity (Subramanian and Feng, 2007; Steidinger *et al.*, 2011); however, some ANG mutants are cytotoxic towards motor neurons, leading to their degradation. A large number of ANG mutants have been identified to be associated with both familial and sporadic amyotrophic lateral sclerosis (ALS) (Conforti *et al.*, 2008; Corrado *et al.*, 2007; Fernández-Santiago *et al.*, 2009; Gellera *et al.*, 2008; Greenway *et al.*, 2006, 2004; Luigetti *et al.*, 2011; Paubel *et al.*, 2008; Seilhean *et al.*, 2009;

Subramanian and Feng, 2007; Subramanian *et al.*, 2008; Thiyagarajan *et al.*, 2012; van Es *et al.*, 2009; Wu *et al.*, 2007; Zou *et al.*, 2012), as well as Parkinson's disease (Rayaprolu *et al.*, 2012; van Es *et al.*, 2011).

2.3.1.5 RNase 6 (RNase k6), RNase 7, and RNase 8

RNase 6 was the sixth characterised member of the human RNase family. It was initially named as RNase k6 to indicate its relationship to the orthologous RNase k2 from bovine kidney. Despite its close sequence similarity to EDN and ECP, RNase 6 is expressed in most tissues, including lung, heart, brain, placenta, liver, kidney, pancreas, and skeletal muscle, but not in eosinophils (Rosenberg and Dyer, 1996). RNase 6 is likely to have a role in host defence due to the presence of RNase 6 encoding mRNA transcripts in neutrophils and monocytes (Sorrentino, 2010); however, no detailed information about this protein is presently available.

The high amino acid sequence identity between RNase 7 and RNase 8 suggests that their genes are likely to have originated from a relatively recent gene duplication event (Sorrentino, 2010). However, the differences in the types of tissues expressing RNase 7 and RNase 8 may suggest that they possess different physiological functions. RNase 7 is found in epithelial tissues and several organs (Harder and Schroder, 2002; Zhang *et al.*, 2003), while RNase 8 is uniquely expressed in the placenta (Zhang *et al.*, 2002). Combining the fact that they both show bactericidal activity against several pathogenic bacteria, both Gram-positive and Gram-negative, the role of RNase 7 possibly lies in the antimicrobial defence of skin (Huang *et al.*, 2007; Abtin *et al.*, 2009; Torrent *et al.*, 2010) while the function of RNase 8 is likely to be protecting the placenta from infections (Rudolph *et al.*, 2006).

2.3.2 Other vertebrate RNases

Besides human RNases, homologues of RNase A have been discovered in many other vertebrate species, including mammals (for instance rodents) (Zhou and Strydom, 1993; Gupta *et al.*, 1999; Dyer *et al.*, 2004), birds (Cho *et al.*, 2005; Nitto *et al.*, 2006), reptiles (Zhao *et al.*, 1994; Nitto *et al.*, 2005), amphibians (Ardelt *et al.*, 1991; Liao *et al.*, 2000), and fish (Pizzo *et al.*, 2006; Cho and Zhang, 2007; Kazakou *et al.*, 2008; Pizzo *et al.*, 2008; Pizzo *et al.*, 2011). Despite their fundamental ribonucleolytic activity, some of these RNases possess more specialised functions, such as the anti-tumour activities of onconase (ONC) from *Rana pipiens* and RC-RNase from *Rana catesbeiana*, and the roles in host defence played by rodent eosinophil-associated ribonucleases (EARs) (Hooper *et al.*, 2003; Ishihara *et al.*, 2003; Garvey *et al.*, 2005), chicken RNase A-2, zebrafish and salmon RNases (Pizzo *et al.*, 2006; Pizzo *et al.*, 2008) to name but a few. Although of therapeutic value, these RNases are not the focus of this study, thus will not be discussed in more detail.

2.4 Inhibitors of RNases

Regulated RNase activity is crucial to organisms; uncontrolled RNA degradation might become cytotoxic and lead to cell death (Vescia *et al.*, 1980; Wu *et al.*, 1993; Leland *et al.*, 1998; Suzuki *et al.*, 1999). Research has revealed both natural and small molecule inhibitors that are capable of inhibiting RNases.

2.4.1 The natural inhibitor of RNases – ribonuclease inhibitor (RI)

The activities of members of the RNase A superfamily can be abolished by the cytosolic ribonuclease inhibitor (RI) (Beintema, Schüller, *et al.*, 1988) (Figure 2.13). RI is a 50 kDa protein that exists in the cytoplasm of all mammalian cells. The functions of RI can be categorised into three classes: ① the regulation of cytoplasmic RNA level; ② the prevention of secreted, non-cytoplasmic RNases from attacking the cytoplasm; and ③ the regulation of the activities of RNases

known to have physiological roles other than the degradation of RNA in the digestive system (Hofsteenge, 1994).

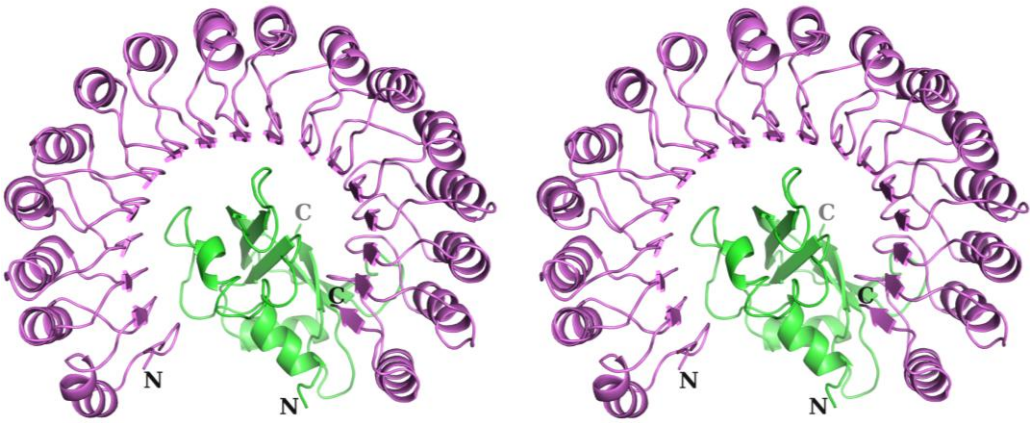


Figure 2.13. Stereo view of the structure of porcine RI (pRI, coloured purple) in complex with RNase A (coloured green) (PDB: 1DFJ; Kobe and Deisenhofer, 1995). The N- and C-termini of both proteins are labelled as N and C respectively.

The inhibition of RNases by RI is achieved via a two-step mechanism, where a rapid formation of an encounter enzyme-inhibitor complex (EI) occurs initially, followed by a slower isomerisation of EI to a tight complex EI^* (Lee *et al.*, 1989; Lee and Vallee, 1989):

$$E + I \leftrightarrow EI \leftrightarrow EI^*$$

RI forms a very tight but reversible, one-to-one complex with RNase. Dissociation constants (K_i) for RI complexed with RNase A (Lee *et al.*, 1989a; Vicentini *et al.*, 1990), ANG (Lee *et al.*, 1989a, 1989b), EDN (Shapiro and Vallee, 1991) are all at femtomolar (fM) level, showing them to be amongst the strongest protein-protein interactions known. More details of RI will be presented in Chapter 4.

2.4.2 Small molecule inhibitors of RNase A

Early studies on the enzymatic mechanism of RNase A suggested that certain chemicals could inhibit the activity of RNase A (Barnard and Stein, 1960; Gundlach *et al.*, 1959; Weil and Seibles, 1955). This led to investigations into

the possibility of using small molecules as RNase A inhibitors. Since the members of the RNase A family recognise their ligands all in a similar way, small molecule inhibitors of RNase A are likely to inhibit its homologues, some of which (for example RNase 4 and ANG) may be involved in disease conditions.

Many such molecules have been characterised to date, such as haloacetates (Barnard and Stein, 1960; Gundlach *et al.*, 1959), aribinonucleotides (Pollard and Nagyvary, 1973), ordinary nucleotides (Richards and Wyckoff, 1971; Eftink and Biltonen, 1987; Shapiro *et al.*, 1989; D'Alessio *et al.*, 1991; Sorrentino *et al.*, 1992; Gilliland, 1997; Shapiro, 2001; Tsirkone *et al.*, 2009) (Figure 2.14), nucleotide analogues (Shapiro *et al.*, 1989; Noguès *et al.*, 1998; Raines, 1998; Parmenopoulou *et al.*, 2012) (Figure 2.15) and derivatives (Tarragona-Fiol *et al.*, 1993; Gilliland, 1997; Leonidas *et al.*, 1997; Leonidas *et al.*, 1999b; Jardine *et al.*, 2001; Shapiro, 2001). RNA based inhibitor design for RNase A has been the most common method used in this field (Russo *et al.*, 1997; Russo and Shapiro, 1999; Russo *et al.*, 2001).

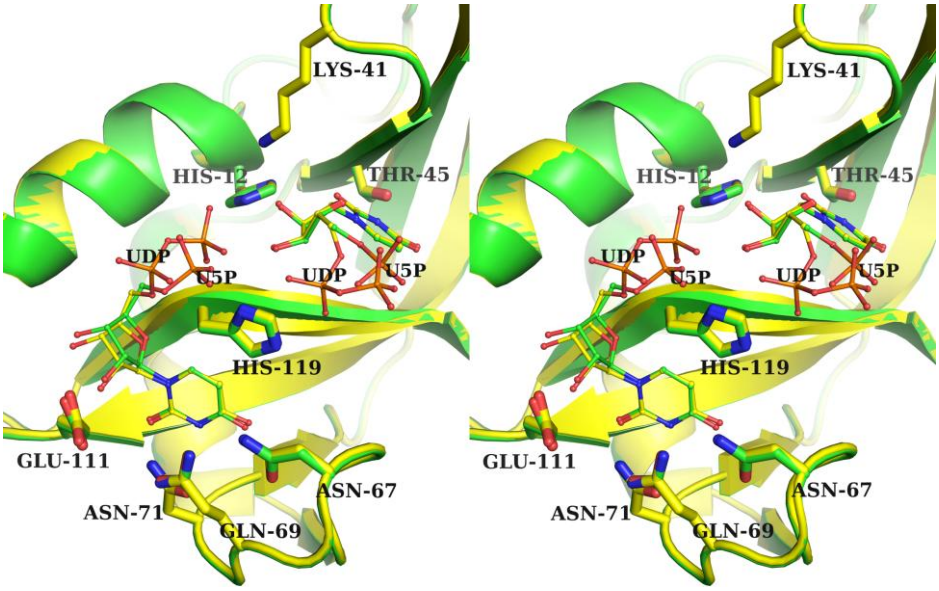


Figure 2.14. Structures of RNase A in complex with uridine 5'-monophosphate (U5P, coloured green), and uridine 5'-diphosphate (UDP, coloured yellow) in stereo view (PDB: 3DXG, 3DXH; Tsirkone *et al.*, 2009). Residues interacting with U5P and UDP are coloured green and yellow respectively. Lys-41 is involved in the binding of UDP, which was not observed in the RNase A-U5P complex structure.

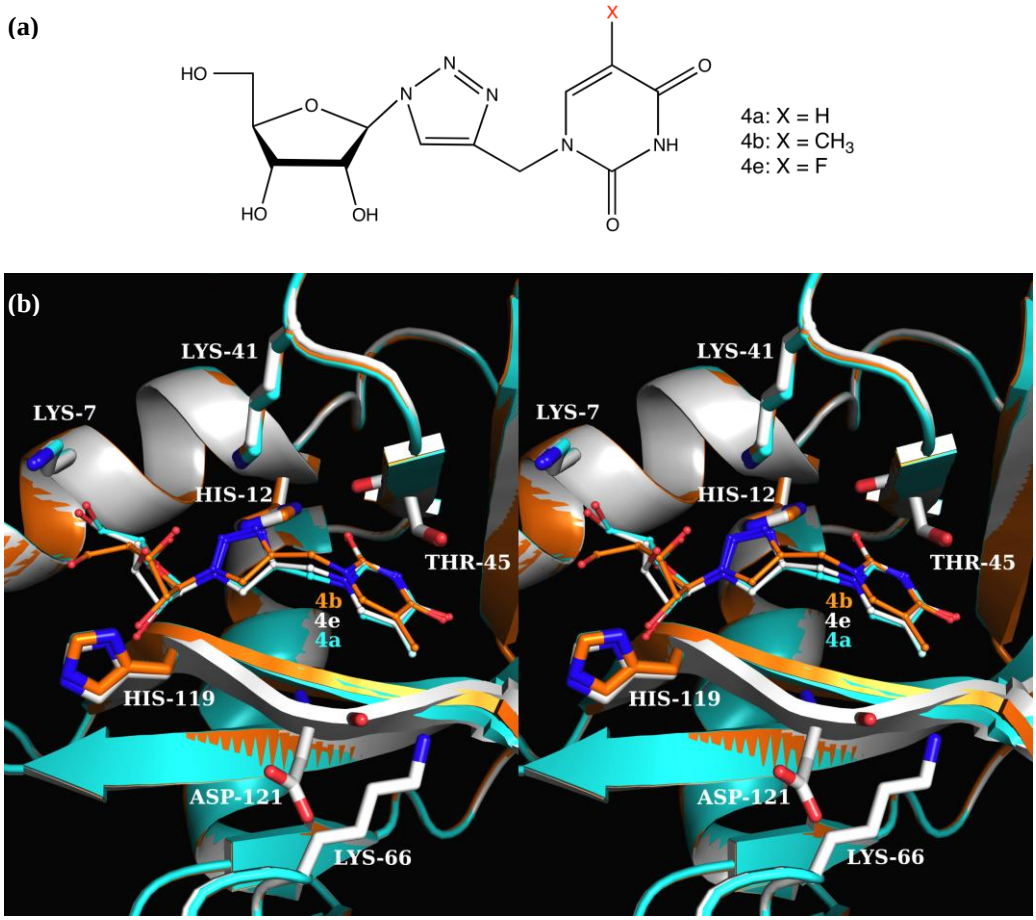


Figure 2.15. RNase A inhibition by triazole pyrimidine nucleosides. (a) Chemical structure of triazole pyrimidine nucleoside analogues; (b) Structures of RNase A in complex with triazole pyrimidine nucleosides 4a (cyan), 4b (orange), and 4e (grey) in stereo view (PDB: 4G8V, 4G8Y, 4G90; Parmenopoulou *et al.*, 2012). Residues interacting with 4a, 4b, and 4e are coloured cyan, orange, and grey respectively. Interactions between inhibitor molecules and His-12, Lys-41, Thr-45, and His-119 are observed in all three structures; Lys-7 interacts with 4a and 4e; Lys-66 and Asp-121 are only involved in the inhibition by 4e.

Chapter 3. Molecular recognition of wild-type and mutant porcine ribonuclease 4 (PL_3) by their mononucleotide substrates

3.1 Introduction

3.1.1 RNase 4, a highly conserved member of the RNase A superfamily

The discoveries of members of the RNase 4 family started with the isolation and characterisation of human RNase 4 (hRNase 4) rom the supernatant of human colon carcinoma conditioned medium (Fett *et al.*, 1985; Shapiro *et al.*, 1986). Three years later, another RNase 4 was identified from porcine liver during the search for RNases that interact with porcine ribonuclease inhibitor (pRI), and was named PL_3 (Hofsteenge *et al.*, 1989). To date, the primary structure of RNase 4 from pig (Hofsteenge *et al.*, 1989), cow (Hosoya *et al.*, 1990), human (Zhou and Strydom, 1993), and rat (Zhao *et al.*, 1998) has been determined by protein sequencing, and that of mouse, chimpanzee, and Sumatran orangutan RNase 4 were deduced from their complementary DNA (cDNA) sequences.

```
rat             QDRMYQRFLRQHVDPEGTGGSDNYCNVMMQRRRMTSTQCKRFNTFIHEDIWNIRSICDTA
mouse           QDRMYQRFLRQHVDPQATGGNDNYCNVMMQRRKMTSVQCKRFNTFIHEDIWNIRGICSTT
Sum_orangutan   QDGMYQRFLRQHVHPEETGGNDRYCNLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICSTT
human           QDGMYQRFLRQHVHPEETGGSDRYCNLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICSTT
chimpanzee      QDGMYQRFLRQHVHPEETGGSDRYCNLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICSTT
porcine         QDRMYQRFLRQHVDPDATGGNDAYCNLMMQRRKMTSHYCKRFNTFIHEDIWNIRSICSTS
bovine          QDGMYQRFLRQHVDPDETGGNDSYCNLMMQRRKMTSHQCKRFNTFIHEDLWNIRSICSTT
                ** **********.*: ***.* ***:*****:**   ***********:****.**.*:

rat             NIPCKNGNMNCHEGIVRVTDCRETGSSVPHNCRYRARASTRRVVIACEGTPEVPVHFDR
mouse           NILCKNGQMNCHEGVVKVTDCRETGNSKAPNCRYRARTSTRRVVIACEGDPEVPVHFDR
Sum_orangutan   NIQCKNGKTNCHEGVVKVTDCRDTGSSRAPNCRYRAMASTRRVVIACEGNPQVPVHFDG
human           NIQCKNGKMNCHEGVVKVTDCRDTGSSRAPNCRYRAIASTRRVVIACEGNPQVPVHFDG
chimpanzee      NIQCKNGKMNCHEGVVKVTDCRDTGSSRAPNCRYRAMASTRRVVIACEGNPQVPVHFDG
porcine         NIQCKNGQMNCHEGVVKVTDCRETGSSRAPNCRYRAMASTRRVVIACEGNPEVPVHFDK
bovine          NIQCKNGQMNCHEGVVRVTDCRETGSSRAPNCRYRAKASTRRVVIACEGNPQVPVHFDG
                ** ****: *****:*:*****:**.*   ****** :*********** *:******
```

Figure 3.1. Comparison of the primary structures of porcine, bovine, mouse, rat, human, Sumatran orangutan, and chimpanzee RNase 4. All sequences were taken from UniProtKB database. The multiple sequence alignment with a conservation scoring was performed using *Clustal Omega* EMBL-EBI online server (McWilliam *et al.*, 2013; Sievers *et al.*, 2011). Conserved residues are highlighted yellow.

Comparison of these primary structures revealed high interspecies similarity (~90%) (Figure 3.1), consistent with the early sequence analysis of porcine, bovine, human and rat RNase 4 (Hofsteenge *et al.*, 1998b). The sequences of

RNase 4 differ only in 30 positions, 18 of which are conservative substitutions. The eight cysteine residues involved in disulphide bridging, the catalytic residues Gln-11, His-12, Lys-41, and His-119 (RNase A numbering), and several residues important for substrate binding are also conserved. Such high level of conservation suggests a strong evolutionary pressure to maintain its structure, which may in turn imply a more specialised function than RNA degradation in the digestive system (Zhou and Strydom, 1993).

Molecular biology studies revealed that the messenger RNA (mRNA) of human RNase 4 (hRNase 4) is found in most tissues, liver being the most abundant source, but not brain (Rosenberg and Dyer, 1995; Futami *et al.*, 1997). In porcine tissues, RNase 4 mRNA is present at high levels in liver and lung, and is detectable in brain but not in kidney, thymus or testis (Vicentini *et al.*, 1994).

Like all RNases in the RNase A superfamily, RNase 4 only cleaves RNA after pyrimidine bases (Shapiro *et al.*, 1986). In contrast to the prototype RNase A that favours cytidine, RNase 4 has a salient feature of strong uridine preference (Vicentini *et al.*, 1994). A comparison of the k_{cat}/K_m [*] for the cleavage of uridylyl-3',5'-adenosine (UpA) and cytidylyl-3',5'-adenosine (CpA) by RNase A and RNase 4 (Table 3.1) suggested that the uridine specificity of RNase 4 results from a suppression of the activity with cytidine-containing substrates (Hofsteenge *et al.*, 1998b).

Table 3.1. Comparison of substrate specificities of RNase A and RNase 4.

Substrate	k_{cat}/K_m ($M^{-1}s^{-1}$)		
	RNase A	**hRNase 4**	**pRNase 4 (PL_3)**
UpA	3.5×10^6	2.5×10^5	2.5×10^5
CpA	4.5×10^6	6.3×10^2	6.7×10^2

[*] k_{cat}/K_m is a measure of how efficiently an enzyme converts a substrate into product. k_{cat}, the turnover number, is the maximum number of substrate molecules converted to product per enzyme molecule per second; K_m, the Michaelis constant, is the substrate concentration at which the reaction rate is at half-maximum.

Residues 37-45 and 120-124 of RNase A contribute to its pyrimidine binding pocket, therefore these areas of PL_3 were altered by site-directed mutagenesis to investigate the structural basis of the different substrate preferences between RNase A and RNase 4. By substituting residues 36-42 of PL_3 with the corresponding ones of RNase A, Vicentini and colleagues observed a large preferential increase in the catalytic efficiency of the hybrid for cytidine-containing substrates, suggesting the involvement of this region in the uridine-preferring substrate specificity of PL_3 (Vicentini *et al.*, 1994). More site-directed mutagenesis studies of PL_3 suggested residues Asp-80 and Arg-101 also contribute to the uridine preference. Furthermore, the substitution of Asp-80 by alanine completely reversed such preference: the k_{cat}/K_m for UpA decreased 159-fold whereas the k_{cat}/K_m for CpA increased 233-fold (Hofsteenge *et al.*, 1998a).

The determination of the crystal structures of free hRNase 4 and in complex with deoxyuridine phosphate (dUp) with a resolution of 2.1 Å and 2.4 Å respectively, provided some structural insights of such substrate preference (Terzyan *et al.*, 1999) (Figure 3.2). The overall structure of hRNase 4 displays the typical kidney-shaped RNase A fold common to all RNases. The P_1 catalytic site of hRNase 4 is similar to that of RNase A, and the residues of catalytic importance are conserved, namely Gln-11, His-12, Lys-40, and His-116, equivalent to Gln-11, His-12, Lys-41, and His-119 of RNase A. The shortened C-terminus of RNase 4 causes structural rearrangement of the side chain of Arg-101 that penetrates the active site pocket and forms hydrogen bond with the O4 atom of uridine. In contrast, the equivalent residues in all other RNases of known structures are located at the surface and not engaged in substrate recognition. The structure of hRNase 4 also revealed that the presence of the side chain of Phe-42 forces Asp-80 to form a hydrogen bond with Thr-44 that in turn becomes an acceptor for the N3 of uridine.

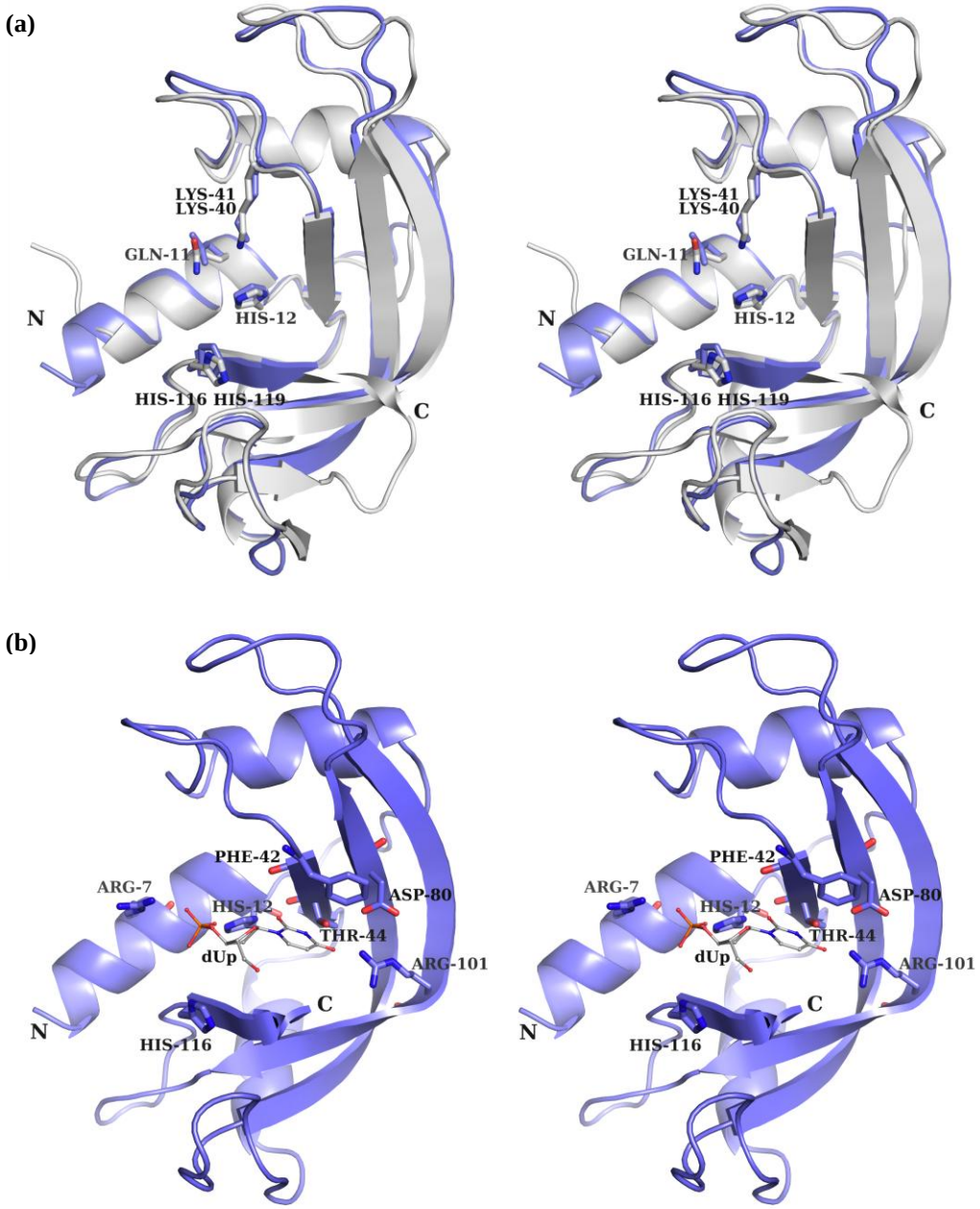


Figure 3.2. Three-dimensional structures of free hRNase 4 and hRNase 4 dUp complex. (a) Superposition of the structures of hRNase 4 (light purple; PDB: 1RNF; Terzyan *et al.*, 1999) and RNase A (grey; PDB: 1AFU; Leonidas et al., 1997). Residues in P_1 subsites of both proteins are shown as sticks; (b) structure of hRNase 4 dUp complex (PDB: 2RNF; Terzyan *et al.*, 1999). Residues Arg-7, His-12, Thr-44, Arg-101 and His-116 interact with dUp ligand via hydrogen bonding, Phe-42 forces Asp-80 to hydrogen bond with Thr-44, which in turn interacts with dUp.

Although with significantly different substrate specificity and possible novel and specialised biological roles, RNase 4 has not been the object of intensive

studies. Only a few reports on the biological functions of RNase 4 have been published so far. Studies revealed that hRNase 4 is a component of the cytoplasmic granules of human peripheral blood monocytes, and this subcellular localisation of RNase 4 suggests that it may be involved in the degradation of polymeric RNA secreted by pathogens, as well as a role that is unique to monocyte-macrophage linkage (Egesten *et al.*, 1997). Recent research suggested that RNase 4 is one of the two RNases secreted by $CD8^+$ T cells that is involved in the inhibition of the X4 strains of HIV-1 (Cocchi *et al.*, 2012). Besides its roles in host defence, hRNase 4 also showed angiogenic activities in *in vitro*, *ex vivo*, and *in vivo* assays, which requires a certain degree of ribonucleolytic activity. More surprisingly, hRNase 4 also showed neurogenic and neuroprotective activities, suggesting its involvement in neurodegeneration diseases such as amyotrophic lateral sclerosis (ALS) (Li *et al.*, 2013).

3.1.2 Rationale behind this study

As stated, biochemical studies suggested that a single amino acid substitution of Asp-80, a highly conserved residue in not only the RNase 4 family but also all the orthologous RNase 1 and some other RNases, to alanine in pRNase 4 completely reversed the substrate specificity from uridine-specific to cytidine-specific. Although the crystal structures of unliganded hRNase 4 and hRNase 4 complexed with d(Up) provided the explanation of the role Asp-80 plays in RNase 4 substrate specificity (Figure 3.2), the initial studies were carried out using porcine RNase 4 – PL_3, whose crystal structure was yet to be solved. The focus of this study was the determination of four crystal structures: ① wild type PL_3 ($wtPL_3$), ② PL_3 D80A mutant, ③ $wtPL_3$ in complex with a uridine-containing substrate, and ④ PL_3 D80A mutant in complex with a cytidine-containing substrate, hoping to provide more structural insights into the role Asp-80 plays in RNase 4 substrate recognition.

3.2 Materials and Methods

3.2.1 DNA sequencing of pYSBLIC-PL_3 gene

The pYSBLIC-PL_3 construct was prepared by K. Kazakou, a previous PhD student in our laboratory, using the pET-YSBLIC vector to carry the PL_3 gene and the *E. coli* BL21-CodonPlus (DE3)-RIL (BL21C$^+$(DE3)-RIL) competent cells as the host (Kazakou, 2008). The pET-YSBLIC vector was a kind gift from Dr M. J. Fogg, University of York, UK. pET-YSBLIC vector is a pET-28a(+) derivative that adds a noncleavable hexahistidine tag directly to the N-terminal of the recombinant protein.

DNA sequencing of the pYSBLIC-PL_3 gene was performed to confirm the identity of the protein. The plasmid of pYSBLIC-PL_3 was extracted from 10 ml of overnight culture using the Wizard plus SV Minipreps DNA purification system (Promega, USA); extraction protocol described in the Wizard plus SV Minipreps DNA purification system Technical Bulletin was followed. The plasmid DNA was sequenced by Eurofins MWG Operon, Germany.

3.2.2 Expression and purification of pYSBLIC-PL_3

While waiting for the DNA sequencing result, an expression and purification trial of pYSBLIC-PL_3 was carried out following the protocols described by Kazakou (2008). 1 L Terrific Broth (TB) medium (1.2% tryptone, 2.4% yeast extract, 72 mM K_2HPO_4, 17 mM KH_2PO_4 and 0.4% glycerol) supplemented with 50 μg/ml kanamycin, and 25 μg/ml chloramphenicol was inoculated with 1 % overnight culture of BL21C$^+$(DE3)-RIL-pYSBLIC-PL_3 cells. The culture was incubated at 37 ℃ with shaking at 200 rpm, and subsequently treated with 1 mM IPTG to induce protein expression when the cells were in their mid-exponential phase of growth (indicated by $OD_{600} \geq 0.6$). The induced cells were further incubated at 37 ℃ with shaking at 200 rpm for 3 hours, and harvested

by centrifugation at 15, 000 g (Beckman JLA 10.500 rotor) for 10 minutes. Cell pellets from 0.5 L culture were stored at -80 ℃.

The cell pellet from 0.5 L bacterial culture was resuspended in 50 ml lysis buffer (50 mM Tris-HCl, pH 8.0) and ruptured using the TS Series Benchtop cell disruptor (Constant Systems Ltd., Daventry, Northants, UK) at 24,000 psi. The lysate was centrifuged at 39,000 g (Beckman JLA 25.50 rotor) for 20 minutes to harvest inclusion bodies (Kazakou, 2008).

The inclusion body pellet was solubilised in 10 ml solubilisation buffer (7 M guanidine hydrochloride, 0.15 M reduced glutathione, 0.1 M Tris-HCl pH 8.0, 2 mM EDTA) for 2 hours. The solubilised protein was separated from the solid matter by centrifugation at 70, 000 g for 30 minutes (Beckman JLA 25.50 rotor), and added dropwisely into 1 L refolding buffer (0.5 M L-arginine, 0.6 mM oxidized glutathione) with gentle stirring. After the addition of the protein sample, the refolding solution was left undisturbed at 4 ℃ overnight (Kazakou, 2008).

pYSBLIC-PL_3 was purified by cation exchange chromatography. A self-packed column with SP-Sepharose Fast Flow beads (GE Healthcare, UK) was equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl). Precipitated protein was removed from the refolding solution by centrifugation at 15, 000 g for 30 minutes (Beckman JLA 10.500 rotor); soluble protein sample was diluted to 2.5 L with Milli-Q H_2O and loaded onto the column. Subsequently the column was re-equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl). pYSBLIC-PL_3 was eluted with an ascending NaCl gradient in equilibration buffer (from 0.1 M to 0.8 M NaCl). Fractions with raised UV trace were analysed by gel electrophoresis (Kazakou, 2008).

The protein sample was further purified using reverse phase chromatography. A Jupiter C4 column (Phenomenex, USA) was used. The column was washed with 95% buffer B (HPLC-grade acetonitrile with 0.08% trifluoroacetic acid (TFA)) and equilibrated with 95% buffer A (degassed HPLC-grade H_2O with 0.1% TFA) prior to protein loading. The column was re-equilibrated with buffer A after the protein was loaded; and the protein was eluted with an increasing Buffer B concentration. Fractions with raised UV trace were pooled and lyophilised overnight. The lyophilised protein sample was dissolved in HPLC-grade H_2O, and analysed by gel electrophoresis (Kazakou, 2008).

3.2.3 Construction of BL21 Codon Plus-pET-46 Ek/LIC-PL_3 strain

The DNA sequencing of the pYSBLIC-PL_3 gene indicated that the construct contained both the wild type PL_3 and a mutant form of PL_3 (details are provided in Section 3.3.1), therefore a new wild type PL_3 construct was prepared using pET-46 Ek/LIC vector from Novagen, Germany.

3.2.3.1 Primer Design

Forward and reverse primers were designed to amplify the gene encoding for PL_3 along with the overhanging bases for ligation-independent cloning (LIC), the codon at position 1 was corrected to encode amino acid residue glutamine, which was the first amino acid of the wild type PL_3 protein (Table 3.2). The primers were synthesised by Invitrogen, USA. The melting temperatures were calculated using equation provided by Invitrogen:

$$TM_{0.05M\ salt} = 59.9 + 41 \times \%GC - \frac{675}{primer\ length}$$

Table 3.2 Primers for amplification of PL_3 gene suitable for LIC cloning.

Forward primer	5'-GAC GAC GAC AAG ATG CAG GAT CGC ATG TAC CAA CGA TT-3'	TM = 62.6 ℃ %GC = 0.50
Reverse primer	5'-GA GGA GAA GCC CGG TTA CTA TTT GTC AAA GTG CAC AGG CAC CT-3'	TM = 65.1 ℃ %GC =0.51

3.2.3.2 Amplification of PL_3 gene

The plasmid extracted from the overnight culture of BL21C$^+$(DE3)-RIL-pYSBLIC-PL_3 was used as the DNA template for the PCR reaction. The primers were dissolved in HPLC-grade H_2O to produce 10 μM stock solutions. The PCR reactions were prepared as shown in Table 3.3. A negative control with no DNA present was carried out. The effect of betaine on DNA amplification was tested. The PCR reaction was performed following the parameters shown in Table 3.4.

Table 3.3. PCR reaction system for the amplification of PL_3 gene.

Forward primer (10 μM)	1 μl
Reverse primer (10 μM)	1 μl
Template PL_3 DNA	0.5 μl
dNTP mix (2 mM each) (Novagen)	5 μl
25 mM $MgSO_4$ (Novagen)	2 μl
5 M betaine	0 or 10 μl
10x PCR Buffer for KOD Hot Start DNA Polymerase (Novagen)	5 μl
KOD Hot Start DNA Polymerase (Novagen)	1 μl
H_2O	x μl
TOTAL	**50 μl**

Table 3.4. Cycling parameters for the amplification of PL_3 gene.

Segment	Cycle(s)	Temperature	Time
1	1	94 ℃	1 minute
2	2-18	94 ℃	1 minute
		58 ℃	30 seconds
		72 ℃	2 minutes
3	1	72 ℃	10 minutes
4	1	4 ℃	-

The PCR product was treated with 10 units DpnI restriction enzyme (Promega) at room temperature for 16 hours to remove the methylated template DNA. The DpnI treated PCR product was analysed by gel electrophoresis. The 400 bp band was cut off from the 1% agarose gel; the recovery of the PCR product was

performed following instructions from the Wizard SV gel and PCR Clean-up system (Promega).

3.2.3.3 T4 Polymerase Treatment of target insert

The PCR product was treated with T4 polymerase (Novagen) to produce sticky ends specifically for LIC cloning. The reactions were prepared as shown in Table 3.5; a negative and a positive control were used. The presence of dATP enabled the 3' to 5' exonuclease activity of T4 DNA polymerase, and produced specific single-strand overhangs. After 30 minutes incubation at 22 ℃, the reactions were stopped by inactivating T4 DNA polymerase at 75 ℃ for 20 minutes.

Table 3.5. Reaction systems for T4 polymerase treatment.

	Negative control	**Positive control**	**PL_3 gene**
T4 DNA polymerase	0.4 µl	0.4 µl	0.4 µl
10× T4 DNA polymerase buffer	2 µl	2 µl	2 µl
25 mM dATP	2 µl	2 µl	2 µl
100 mM DTT	1 µl	1 µl	1 µl
Control insert/PCR product	0 µl	4 µl	0.2 pmol*
H_2O	14.6 µl	10.6 µl	Add H_2O to 20 µl

* 0.2 pmol of PCR product was required in the reaction, the volume of which can be calculated based on equation: $volume\ of\ DNA = \frac{the\ number\ of\ moles\ in\ pmol \times 0.66 \times DNA\ size\ in\ bp}{DNA\ concentration\ in\ \mu g/ml}$

3.2.3.4 Annealing of the insert and pET-46 Ek/LIC vector

The T4 polymerase treated insert was subsequently used to prepare the pET-46 Ek/LIC-PL_3 plasmid. 1 µl of pET-46 Ek/LIC vector from the pET-46 Ek/LIC vector kit (Novagen) and 2 µl T4 DNA polymerase treated PL_3 gene insert (0.02 pmol) were mixed and incubated at 22 ℃ for 5 minutes. Subsequently, 1 µl of 25 mM EDTA was added and incubated for further 5 minutes at 22 ℃.

3.2.3.5 Preparation of pET-46 Ek/LIC-PL_3 cloning strain

The annealed pET-46 Ek/LIC-PL_3 plasmid was used to prepare the plasmid-amplifying strain. Top10 *E. coli* cells (Invitrogen) were used as the host. Top10 cells inoculated with 2 µl pET-46 Ek/LIC-PL_3 plasmid were incubated on ice for 5 minutes and heat-shocked at 42 ℃ for 90 seconds. The heat-shocked cells were kept on ice for 2 minutes before incubating with 500 µl LB broth at 37 ℃ for 1 hour. The cells were plated on an LB agar plate supplemented with 100 µg/ml ampicillin and incubated at 37 ℃ with shaking at 200 rpm overnight. 10 ml LB broth supplemented with 100 µg/ml ampicillin was inoculated with an isolated colony from the plate and incubated at 37 ℃ overnight. pET-46 Ek/LIC-PL_3 plasmid DNA was extracted from the cells using the Wizard plus SV Minipreps DNA purification system (Promega); the plasmid DNA was sequenced by Eurofins MWG Operon.

3.2.3.6 Preparation of pET-46 Ek/LIC-PL_3 expression strain

The recombinant plasmid DNA was used to prepare the expression strain. BL21-CodonPlus (DE3)-RIPL (BL21C$^+$(DE3)-RIPL) *E. coli* cells (Agilent Technologies, Santa Clara, California, USA) were used as the host; the same transformation protocol described in Section 3.2.3.5 was followed. A glycerol stock of this construct was prepared by mixing 200 µl of the overnight culture with 800 µl autoclaved 50% glycerol. The stock was stored at -80 ℃ after freezing in liquid N_2.

3.2.4 Expression and purification of wild type pET46-PL_3 (wtPL_3)

The expression and purification protocols detailed in Section 3.2.2 were followed to express and purify wtPL_3 from BL21C$^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 strain. Modifications were made to the protocols to optimise protein yield (the incubation time after induction was extended from 3 hours to 16 hours; the volume of refolding buffer was reduced from 1 L to 500 ml),

which will be discussed in the Results section. Optimised expression and purification protocols were given in Appendix II.

3.2.5 Mass spectrometry analysis of tagged wild-type pET46-PL_3

The identity of the tagged PL_3 produced was verified by mass spectrometry. The experiment was carried out using an Agilent 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) mass spectrometer (Agilent Technologies). The protein sample was in HPLC-grade H_2O at a concentration of 10 mg/ml. Data acquisition and processing was controlled via Agilent MassHunter Workstation Software (Agilent Technologies). The mass spectrometry experiment was performed by Mr M. Lewis from the Department of Chemistry, University of Bath.

3.2.6 His-tag cleavage of wtPL_3 by Enterokinase

3.2.6.1 Pilot study of His-tag cleavage of wtPL_3 by enterokinase

EnterokinaseMax (EKMax) from Invitrogen and recombinant enterokinase (rEk) from New England BioLabs (NEB), USA were used in the pilot study of PL_3 tag cleavage. The preparation of the pilot reactions were based on the instruction described in the EKMax User Manual (Invitrogen).

The reactions contained 20 µg PL_3 protein, 3 µl 10× EKMax Buffer (500 mM Tris-HCl, pH 6.5, 10 mM $CaCl_2$, 1% Tween-20 (v/v)), and six different amounts of EKMax (0 unit[†] as negative control, 0.001 unit, 0.01 unit, 0.1 unit, 1 unit, and 4 units) (Table 3.6). 1× EKMax Buffer was used for the negative control, and to prepare the 10-fold serial dilutions of EKMax; 4 µl of undiluted EKMax was used for the 4 units of EKMax. The rEk was diluted with 1×

[†]1 unit of EKMax is the amount of enzyme required to digest 20 µg of a thioredoxin-CAT fusion protein to 90% completion in 16 hours at 37 ℃.

EKMax Buffer to prepare an enzyme solution with a concentration of 0.00016 µg/µl, which is equivalent to 1 unit/µl according to the unit definition provided in the datasheet.

Table 3.6. Small-scale optimisation for enterokinase treatment.

PL_3 protein	20 µg
10× EKMax Buffer	3 µl
EKMax (or 1× EKMax Buffer for the negative control)	1 µl or 4 µl
Deionized H_2O	x µl
Total volume	**30 µl**

PL_3 protein used in this study was purified by ion exchange chromatography followed by reverse phase chromatography, lyophilised and re-suspended in HPLC-grade H_2O. The concentration of the PL_3 sample was measured by NanoDrop 2000c spectrophotometer (Thermo Scientific, USA) using the entered Abs 1%[‡] = 7.69 $(g/L)^{-1}mm^{-1}$. The reactions were incubated at room temperature for 16 hours. The efficiency of the enterokinase cleavage was analysed by SDS-PAGE.

3.2.6.2 Scale-up of Enterokinase cleavage reaction

The pilot study showed that EKMax and rEk had similar levels of activity, and 1 unit of enterokinase gave complete cleavage of 20 µg PL_3 protein in 1× EKMax Reaction Buffer after 16 hours at room temperature.

For large-scale enterokinase cleavage reactions, rEK was used; 10× EKMax Reaction Buffer (500 mM Tris-HCl, pH 6.5, 10 mM $CaCl_2$, 1% Tween-20 (v/v)) was prepared as instructed in the EKMax User Manual. The reaction system using 1 unit of enterokinase was scaled up in a linear manner to suit the amount

[‡] Abs 1% is calculated using the equation: $Abs\ 1\% = 10 \times \frac{\varepsilon}{Mr}$, where ε is the molar extinction coefficient of the protein; Mr is the molecular weight of the protein.

of PL_3 to be treated. No dilution was made to the PL_3 sample prior to the rEk treatment to achieve better cleavage results.

3.2.7 Purification of untagged wtPL_3 by affinity chromatography

The rEk was removed from the PL_3 sample by affinity chromatography. A 1 ml HiTrap Benzamidine FF (High Sub) column (GE Healthcare) was used to trap the rEk protease; all operations were performed with a syringe and a luer adapter provided. The rEk treated PL_3 sample was loaded onto the column that had been equilibrated with 5 column volumes of equilibration buffer (50 mM Tris-HCl, pH 8.0, 1 mM $CaCl_2$), and the column was subsequently washed with 5 column volumes of equilibration buffer and 5 column volumes of wash buffer 2 (50 mM Tris-HCl, pH 8.0, 100 mM NaCl). Untagged PL_3 did not interact with benzamidine, and would be present in the flow through of the loading and first washing steps. Bound EKMax was eluted from the column with 5 column volumes of elution buffer (50 mM Tris-HCl, pH 8.0, 1 M NaCl). Fractions from loading and first wash step were collected in separate tubes and analysed by SDS-PAGE.

3.2.8 Purification of untagged wtPL_3 by reverse phase chromatography

The untagged PL_3 protein was further purified using reverse phase HPLC chromatography. The purification protocol described in 3.2.2 was followed.

3.2.9 Construction of pET46-PL_3 D80A mutant

3.2.9.1 Primer Design

A set of primers was designed for the site-directed mutagenesis of wild type PL_3 DNA to produce PL_3 D80A mutant (Table 3.7). The primers were manufactured by Eurofins MWG Operon. The melting temperatures were calculated using the equation given in Section 3.2.3.1.

Table 3.7. Primers for amplifying PL_3 D80A mutant by site-directed mutagenesis.

Forward primer	5'- GTG GTG AAG GTC ACA GCG TGC AGG GAG ACA GGA AG -3'	TM = 65.2 ℃ %GC = 0.60
Reverse primer	5'- CTT CCT GTC TCC CTG CAC GCT GTG ACC TTC ACC AC -3'	TM = 65.2 ℃ %GC =0.60

3.2.9.2 PCR Site-directed mutagenesis of wtPL_3 to produce PL_3 D80A DNA

The plasmid DNA extracted from the overnight culture of BL21C$^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 was used as the template DNA. The primers were dissolved in HPLC-grade H_2O to produce 10 μM stock solutions. The PCR recipe in Table 3.2 was followed. 2 μl of PL_3 template DNA was used instead of 0.5 μl. The PCR reactions were carried out following the parameters shown in Table 3.3. Time for annealing and elongation was set to 30 seconds and 6 minutes respectively.

3.2.9.3 Preparation of pET46-PL_3 D80A cloning strain

The PCR product was treated with 10 units DpnI restriction enzyme at room temperature for 2 hours to digest methylated template DNA and subsequently used in the preparation of the cloning strain of pET-46 Ek/LIC-PL_3 D80A. Top10 *E. coli* competent cells were used as the host; the transformation protocol described in Section 3.2.3.5 was followed. pET-46 Ek/LIC-PL_3 D80A plasmid DNA was extracted from the cells using the Wizard plus SV Minipreps DNA purification system (Promega); the plasmid DNA was sequenced by Eurofins MWG Operon.

3.2.9.4 Preparation of pET46-PL_3 D80A expression strain

Once the DNA sequencing result was confirmed correct, the recombinant plasmid DNA of pET46-PL_3 D80A was used to prepare the expression strain. BL21C$^+$(DE3)-RIPL *E. coli* cells (Agilent Technologies) were used as the host;

the same transformation protocol described in Section 3.2.3.6 was followed. A glycerol stock of the BL21C^{+}(DE3)-RIPL-pET-46 Ek/LIC-PL_3 D80A construct was prepared by mixing 200 µl of the overnight culture with 800 µl autoclaved 50% glycerol. The stock was stored at -80 ℃ after freezing in liquid N_2.

3.2.10 Expression and purification of pET46-PL_3 D80A mutant

The optimised expression and purification protocols of wtPL_3 protein (Appendix II) were followed to express and purify the PL_3 D80A mutant.

The N-terminal hexahistidine tag of the PL_3 D80A mutant protein was removed by enterokinase following the protocol described in Section 3.2.6.2.

3.2.11 Mass spectrometry analysis of wtPL_3 and PL_3 D80A mutant samples

Identity of the three PL_3 samples (namely untagged wtPL_3, tagged and untagged PL_3 D80A mutant) was verified by mass spectrometry. The experimental procedure described previously in Section 3.2.5 was followed. All three PL_3 samples were in HPLC-grade H_2O with a concentration of 10 mg/ml.

3.2.12 Ribonucleolytic activity assay of wtPL_3 and PL_3 D80A mutant samples

The RNase assay protocol described by Shapiro *et al.* (1987), which was originally designed to assess the ribonucleolytic activity of hANG, was carried out to test the activities of wtPL_3 and PL_3 D80A mutant samples; human angiogenin (hANG) protein was used as a reference.

The assays contained 10 mg/ml yeast tRNA, 1 mg/ml bovine serum albumin (BSA), and six concentrations of PL_3 samples (0-500 nM) in RNase assay buffer (0.1 M NaCl, 0.1 M HEPES pH 7.0) (Table 3.8). All assays were carried out in duplicate. Reactions were terminated by the addition of 700 µl ice-cold 3.4% perchloric acid after 2 hours incubation in 37 ℃ water bath. The mixtures were spun down at 16,100 g (Eppendorf Microcentrifuge 5415D) for 10 minutes at 4 ℃ and the absorbance at 260 nm (A_{260}) of the supernatants were used as a measure of the ribonucleolytic activity of test proteins. The magnitudes of the A_{260} values in the quasilinear portion of each plot (300-500 nM for hANG, 0.3-0.5 nM for test samples) were used in the calculation of relative activities.

Table 3.8. RNase activity assay reaction system.

Assay mixture	**Concentration of protein sample (µM)**						
	blank	0.05	0.1	0.2	0.3	0.4	0.5
Sep-pak water* (µl)	110	105	100	90	80	70	60
RNase assay buffer (µl)	100	100	100	100	100	100	100
1 mg/ml BSA (µl)	30	30	30	30	30	30	30
40 µg/ml test protein (µl)	0	5	10	20	30	40	50
10 mg/ml tRNA (µl)	60	60	60	60	60	60	60
Total (µl)	300	300	300	300	300	300	300

*Prepared using Sep-Pak Classic cartridges (Waters, USA).

3.2.13 Crystallisation of wtPL_3 and PL_3 D80A mutant

The 4 protein samples listed in Table 3.9 were used in the pre-crystallisation tests (PCTs) and initial screenings of crystallisation conditions.

Table 3.9. Proteins used in the PCT tests and screening of crystallisation conditions.

Protein samples used in PCT tests and screening of crystallisation conditions
tagged wtPL_3
untagged wtPL_3
tagged PL_3 D80A mutant
untagged PL_3 D80A mutant

3.2.13.1 Pre-crystallisation Tests (PCTs)

A PCT test was performed with samples listed in Table 3.9 to establish a starting concentration for the subsequent screenings of crystallisation conditions. The PCT reagents (Table 3.10) were self-prepared in the laboratory, and the experimental protocol described in the PCT User Guide by Hampton Research was followed with minor alterations as detailed below. All samples were diluted with HPLC-grade H_2O to a concentration of 10 mg/ml.

1 ml of PCT reagent A1 was pipetted into the well A1 of a Nunclon multidish 4-well plate. 1 µl of protein sample was pipetted onto the centre of a glass cover slide (Hampton Research), and an equal volume of reagent A1 from the well was pipetted into the protein drop without mixing. The cover slide was inverted over well A1, which was sealed with vacuum grease (Dow Corning®). The procedure was repeated for reagents A2, B1, and B2. The plate was incubated at 16 ℃ for 30 minutes prior to results examination.

Table 3.10. Formulations of PCT reagents.

PCT reagents	Formulation
Reagent A1	0.1 M Tris-HCl pH 8.5, 2.0 M $(NH_4)_2SO_4$
Reagent A2	0.1 M Tris-HCl pH 8.5, 1.0 M $(NH_4)_2SO_4$
Reagent B1	0.1 M Tris-HCl pH 8.5, 0.2 M $MgCl_2 \cdot 6H_2O$, 30% w/v PEG 4000
Reagent B2	0.1 M Tris-HCl pH 8.5, 0.2 M $MgCl_2 \cdot 6H_2O$, 15% w/v PEG 4000

3.2.13.2 Screening for crystallisation conditions for wtPL_3 and PL_3 D80A mutant

Protein samples listed in Table 3.9 (tagged wtPL_3, untagged wtPL_3, tagged PL_3 D80A, and untagged PL_3 D80A) were diluted with HPLC-grade H_2O to 10

mg/ml, which was determined to be an appropriate starting concentration for crystallisation by PCT tests.

Crystallisation screens Structure screen I & II HT-96, PACT premier HT-96, JCSG-plus HT-96, and Morpheus HT-96 (Molecular Dimensions) were used to set up the initial screenings. The screenings were set up in 96-well Intelli crystallisation plates (Art Robbins Instruments, California, USA) by Phoenix protein crystallisation robot (Art Robbins Instruments). Equal volumes (0.15 µl) of protein samples and reservoir solutions were added to the drop wells to give a final protein concentration of 5 mg/ml. All plates were incubated at 16 ℃.

3.2.13.3 Optimisation of crystallisation conditions for wtPL$_3$ and PL$_3$ D80A mutant

Microcrystals of wtPL$_3$ and PL$_3$ D80A mutant were observed in several crystallisation conditions from Structure screen I & II HT-96, JCSG-plus HT-96, and Morpheus HT-96. Crystallisations were repeated using the hanging drop vapour diffusion method in 24-well XRL plates (Molecular Dimensions) with increased or decreased salt and/or precipitant concentrations to optimise protein crystallisation.

3.2.14 Crystallisation of wtPL$_3$ ·dUMP and PL$_3$ D80A ·dCMP complexes

3.2.14.1 Soaking tagged wtPL$_3$ crystals with (UpA)$_5$ and dUMP ligands

Crystallisation of tagged wtPL$_3$ was repeated using condition Morpheus – E10 (0.1 M Morpheus Buffer System 3 pH 8.5, 0.12 M Monosaccharides, 30% v/v EDO_P8K) in 24-well XRL plates by hanging drop vapour diffusion method, with a final drop size of 2 µl. Crystals of tagged wtPL$_3$ took 3 days to grow.

Initial attempt of $wtPL_3$ ligand recognition was carried out using a RNA oligonucleotide $(UpA)_5$ synthesised by Eurofins MWG Operon. The $wtPL_3$ protein crystals were soaked with 2 mM $(UpA)_5$ solution for 5 minutes, 1 hour, and 6 hours at room temperature respectively, flash-frozen with liquid nitrogen using the reservoir solution as cryo-protectant and transported to Diamond Light Source (DLS; Didcot, Oxfordshire, UK) for data collection.

The mononucleotide ligand 2'-deoxyuridine-5'-monophosphate (dUMP) was also used in the soaking experiment. The soaking experiments were performed with two final concentrations of dUMP, 16.7 mM and 33.3 mM respectively, and four concentrations of ethylene glycol ranging from 16.7% to 28.3% as a cryo-protectant. Three sets of soaking experiments were set up 3 days, 2 days, and 1 day prior to crystal harvesting. Soaked $wtPL_3$ crystals were flash-frozen with liquid nitrogen and transported to DLS, UK for data collection.

3.2.14.2 PCT tests for co-crystallisations of PL_3 and their ligands

Due to the fact that soaking $wtPL_3$ crystals with uridine-containing ligand failed to produce complex crystals, and PL_3 D80A crystals grew as clusters that were not suitable for soaking experiments, the method of co-crystallisation was pursued. The ligands used were 2'-deoxyuridine-5'-monophosphate (dUMP) and 2'-deoxycytidine-5'-monophosphate (dCMP).

PCT tests were performed with both the $wtPL_3$ dUMP and PL_3 D80A dCMP to establish a starting concentration for subsequent crystallisation trials; the PCT protocol described in Section 3.2.13.1 was followed. Four protein samples listed in Table 3.9, namely tagged $wtPL_3$, untagged $wtPL_3$, tagged PL_3 D80A, and untagged PL_3 D80A proteins were used. Protein samples of $wtPL_3$ and PL_3 D80A were mixed with dUMP and dCMP respectively in a 1:1.2 molar ratio, and the protein concentrations in the final solution were 9 mg/ml.

3.2.14.3 Screening for crystallisation conditions for wtPL$_3$ dUMP and PL$_3$ D80A dCMP complexes

It was established from the PCT tests that a 9 mg/ml protein sample with 1.2× ligand in molarity was an appropriate starting point for the initial screenings. Crystallisation screens Structure Screen I & II HT-96, PACT premier HT-96, JCSG-plus HT-96, ProPlex HT-96, Stura Footprint + MacroSol HT-96, Morpheus HT-96, and MIDAS HT-96 (Molecular Dimensions) were used to set up the initial screenings. Protocol described in Section 3.2.13.2 was followed.

Additive screenings for PL$_3$ D80A dCMP crystallisation conditions were also carried out. 90% of Structure Screen I & II – A7 was used as the main crystallisation condition, to which 10% Silver Bullets (Hampton Research), Structure screen I & II HT-96, or MIDAS HT-96 was added. The additive screenings were performed using Phoenix protein crystallisation robot (Art Robbins Instruments) following the protocol described in Section 3.2.13.2.

3.2.14.4 Optimisation of crystallisation conditions for wtPL$_3$ dUMP and PL$_3$ D80A dCMP complexes

Microcrystals of wtPL$_3$ dUMP and PL$_3$ D80A dCMP were observed in several crystallisation conditions from Structure Screen I & II HT-96, PACT premier HT-96, JCSG-plus HT-96, and Stura Footprint + MacroSol HT-96. Crystallisations were repeated using the hanging drop vapour diffusion method in 24-well XRL plates (Molecular Dimensions) with increased or decreased salt or precipitant concentrations to optimise protein crystallisation.

3.2.15 Data collection and structure solution of wtPL$_3$, PL$_3$ D80A mutant, wtPL$_3$ dUMP and PL$_3$ D80A dCMP complexes

3.2.15.1 Data collection and structure solution of wtPL$_3$

Tagged wtPL$_3$ protein crystals obtained from condition Morpheus – F10 were flash-frozen with liquid nitrogen using the reservoir solution as cryo-protectant and transported to DLS, UK for data collection. X-ray diffraction data for wtPL$_3$ were collected from a single crystal with 0.1 °oscillation using a Pilatus3 6M detector at beamline I24 with 0.02 second exposure time.

The diffraction data were automatically indexed and integrated with *XDS* (Kabsch, 2010) and *xia2* (Winter *et al.*, 2013), and scaled to 1.25 Å using *AIMLESS* from the *CCP4* programme suite (Evans, 2006, 2011; Winn *et al.*, 2011). Initial phases were obtained by molecular replacement using *Phaser MR* from the *CCP4* programme suite (McCoy *et al.*, 2007; Winn *et al.*, 2011). Structure of human RNase 4 (PDB: 1RNF; Terzyan *et al.*, 1999) was used as a the search model. Isotropic and anisotropic structural refinements (the latter is often used on structures with higher resolutions) were performed using *REFMAC5* from the *CCP4* programme suite (Murshudov *et al.*, 1997; Murshudov *et al.*, 2011; Winn *et al.*, 2011) and *COOT* version 0.7.1 (Emsley *et al.*, 2010). The alternative conformations of residues and ethylene glycol molecules were added manually, and water molecules were added to the model using *COOT* (Emsley *et al.*, 2010) at positions where $mF_o - DF_c$ electron density peaks exceeded 3σ. *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999) and *MolProbity* (Davis *et al.*, 2007; Chen *et al.*, 2010) were used for structure validation. Figures were drawn with *PyMOL* (DeLano, 2002).

3.2.15.2 Data collection and structure solution of PL$_3$ D80A mutant

Tagged PL$_3$ D80A protein crystals obtained from condition JCSG-plus – E1 were flash-frozen by liquid nitrogen with 25% glycerol as cryo-protectant and transported to DLS, UK for data collection. X-ray diffraction data for PL$_3$

D80A were collected from a single crystal with 0.5 ° oscillation using a Pilatus 6M-F detector at beamline I04.

The protocol described in Section 3.2.15.1 was followed to solve and refine the structure of PL_3 D80A mutant. Integrated diffraction data were scaled to 2.9 Å, and the structure of $wtPL_3$ was used as a search model in molecular replacement. Structure refinements were performed isotropically, as the resolution of the dataset was not high enough for anisotropic refinement. Automatically generated local non-crystallographic symmetry (NCS) restraints were used in refinements.

3.2.15.3 Data collection and structure solution of $wtPL_3$ dUMP complex

Untagged $wtPL_3$ dUMP crystals obtained from condition ProPlex – E7 were flash-frozen with liquid nitrogen with 25% glycerol as cryo-protectant and transported to DLS, UK for data collection. X-ray diffraction data for $wtPL_3$ dUMP were collected from a single crystal with 0.1 ° oscillation using a Pilatus 2M detector at beamline I04-1.

The protocol described in Section 3.2.15.1 was followed to solve and refine the structure of the $wtPL_3$ dU5P complex. Integrated diffraction data were scaled to 2.28 Å, and the structure of $wtPL_3$ was used as a search model in molecular replacement. Structure refinements were performed isotropically.

3.2.15.4 Data collection and structure solution of PL_3 D80A dCMP complex

Tagged PL_3 D80A dCMP crystals obtained from modified condition Structure Screen I & II – A9 were flash-frozen with liquid nitrogen with 25% glycerol as cryo-protectant and transported to DLS, UK for data collection. X-ray diffraction data for PL_3 D80A dCMP were collected from a single crystal with 0.5 ° oscillation using a Pilatus 6M detector at beamline I02.

The protocol described in Section 3.2.15.1 was followed to solve and refine the structure of PL_3 D80A dCMP complex. Integrated diffraction data were scaled to 2.6 Å, and the structure of $wtPL_3$ was used as a search model in molecular replacement. Structure refinements were performed isotropically. Automatically generated local non-crystallographic symmetry (NCS) restraints were used in refinements.

3.3 Results

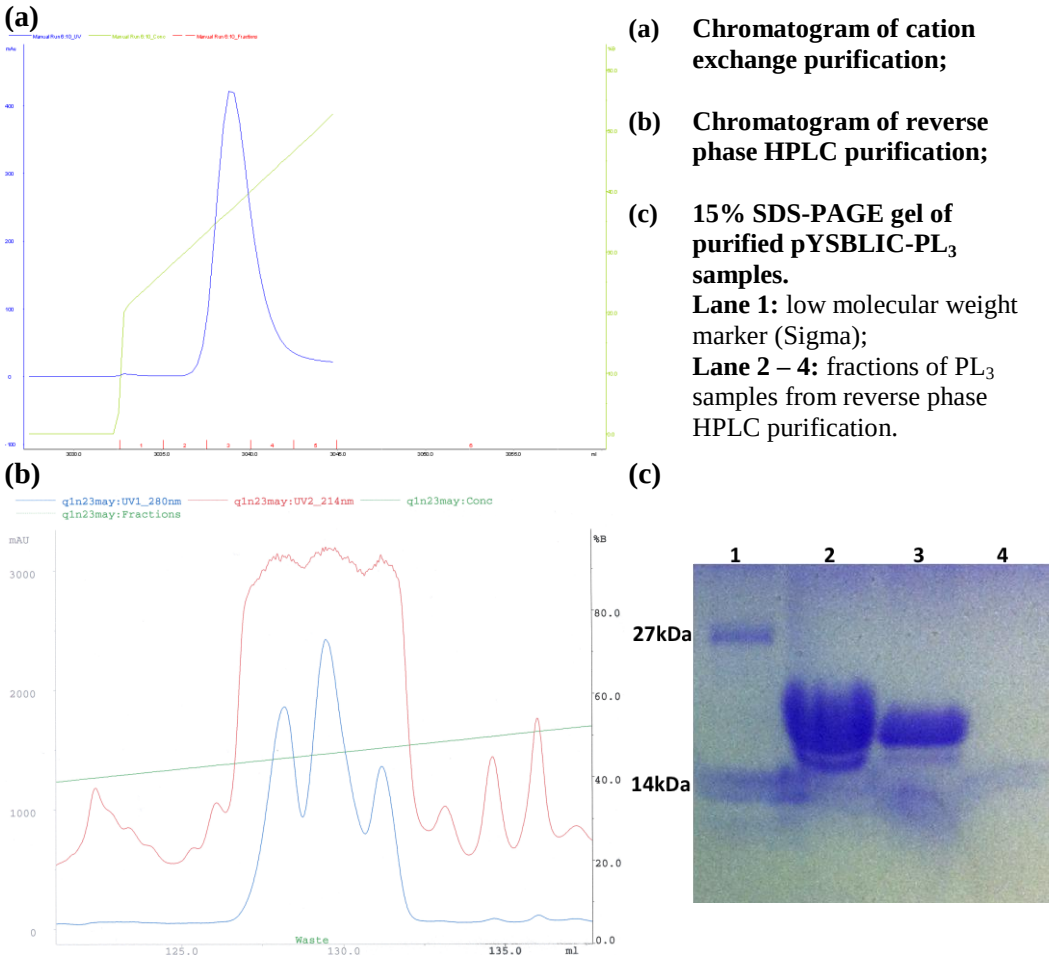
There were two PL_3 constructs available in the laboratory at the time, ptRP-PL_3 and pYSBLIC-PL_3. The ptRP-PL_3 construct was obtained from our collaborator Dr Jan Hofsteenge from Basel, Switzerland. The pYSBLIC-PL_3 construct was prepared by K. Kazakou, a previous member of our laboratory.

ptRP-PL_3 would produce PL_3 protein without N- or C-terminal tags. However, the lack of a detailed vector map of the self-synthesised ptRP vector, as well as the prolonged timespan of expression had made the pYSBLIC-PL_3 construct the ideal choice.

3.3.1 DNA sequencing of pYSBLIC-PL_3 gene, expression and purification of pYSBLIC-PL_3

The plasmid DNA of pYSBLIC-PL_3 was extracted using the Wizard plus SV Minipreps DNA purification system (Promega) and sequenced by Eurofins MWG Operon to confirm the DNA sequence. An expression and purification trial of pYSBLIC-PL_3 was carried out following the expression and purification protocols described in Section 3.2.2.

Pure pYSBLIC-PL_3 could be purified following the purification protocol at a concentration of ~10 mg/ml. However, the DNA sequencing detected a Gln to Asn substitution at position 1 of the mature protein. Such mutation was not detected after the initial construction of this strain. Hence, a new construct for the wild type PL_3 was prepared.



(a) Chromatogram of cation exchange purification;

(b) Chromatogram of reverse phase HPLC purification;

(c) 15% SDS-PAGE gel of purified pYSBLIC-PL$_3$ samples.
Lane 1: low molecular weight marker (Sigma);
Lane 2 – 4: fractions of PL$_3$ samples from reverse phase HPLC purification.

Figure 3.3. Results of pYSBLIC-PL$_3$ purifications. (a) Chromatogram of cation exchange purification of pYSBLIC-PL$_3$; (b) Chromatogram of reverse phase HPLC purification of pYSBLIC-PL$_3$; (c) 15% SDS-PAGE gel of purified pYSBLIC-PL$_3$ samples.

3.3.2 Construction of BL21C$^+$-pET-46 Ek/LIC-PL$_3$

The protocol described in Section 3.2.3 was followed to prepare the BL21C$^+$(DE3)-RIPL-pET-46 EK/LIC-PL$_3$ construct. DpnI was used to remove methylated template DNA, and tThe DpnI treated PCR products were analysed by gel electrophoresis (Figure 3.4). Lane 1 and 2 are the negative controls without and with betaine, which reduces secondary structure formation in GC-rich regions (Rees *et al.*, 1993); lane 3 and 4 are the PL$_3$ PCR products without and with betaine respectively. The 400 bp band in both PCR samples matches the size of the PL$_3$ gene.

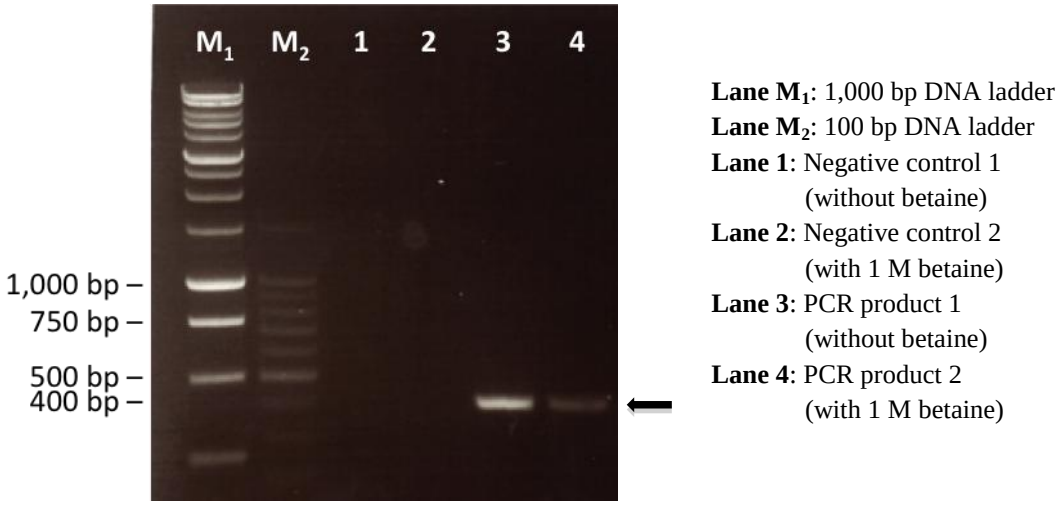


Figure 3.4. DpnI treated PCR products on a 1% agarose gel. Neither of the negative controls contained DNA; both of the PCR products showed a single band of DNA at 400 bp, which matched the size of the PL_3 gene.

The DNA of PL_3 gene was recovered from the gel and inserted into the pET-46 Ek/LIC vector. The pET-46 Ek/LIC-PL_3 plasmids were transformed into, and extracted from the overnight cultures of Top 10 and BL21 Codon Plus (BL21C$^+$(DE3)-RIPL) competent cells respectively. The plasmid DNA was sequenced, and the results were aligned with the wild type PL_3 DNA obtained from the NCBI database (Gene ID: 396976) (Figure 3.5).

The alignment confirmed that the pET-46 Ek/LIC-PL_3 plasmid as well as the BL21C$^+$-pET-46 Ek/LIC-PL_3 expression strain were successfully constructed.

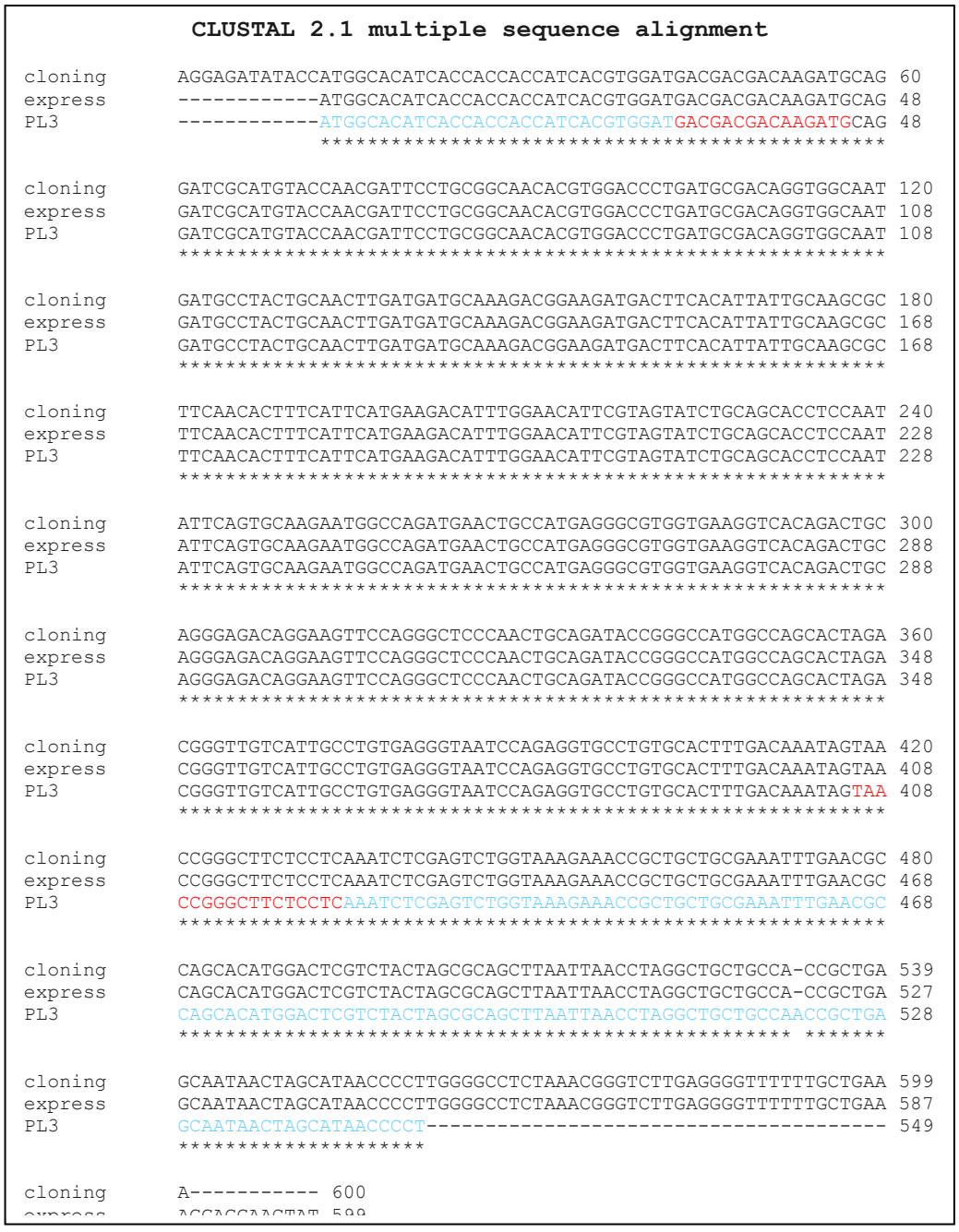


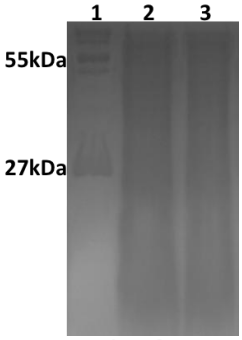
Figure 3.5. Sequence alignments between the sequences of the cloning plasmid, the expression plasmid, and the PL_3 gene sequence obtained from NCBI database (Gene ID: 396976). Alignment was performed using Clustal W2 (Larkin *et al.*, 2007). Sequence coloured in blue and red represent the sequences of pET-46 Ek/LIC vector and the primers used in the PCR reaction respectively. Coloured sequences are not part of the native PL_3 gene; they were included in the alignment to ensure the correctness of the remainder of the target sequence.

3.3.3 Expression and purification of wild-type pET46-PL_3

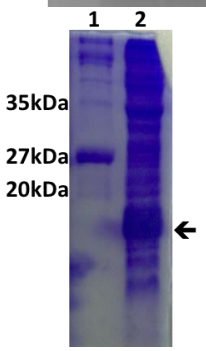
The expression and purification protocol described in Section 3.2.2 was followed to express and purify $wtPL_3$ from the $BL21C^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 strain.

3.3.3.1 Optimisation of the expression condition for pET46-PL_3

No significant expression of PL_3 could be observed 3 hours after IPTG induction (Figure 3.6a). To increase the protein yield, the time for protein expression was extended: induced culture was incubated at 16 ℃ for 16 hours. Protein yield was improved (Figure 3.6b).



(a) Level of PL_3 expression after 3 hour induction at 37 ℃.
Lane 1: low molecular weight marker (Sigma);
Lane 2: cell culture before induction;
Lane 3: cell culture after 3-hour induction.

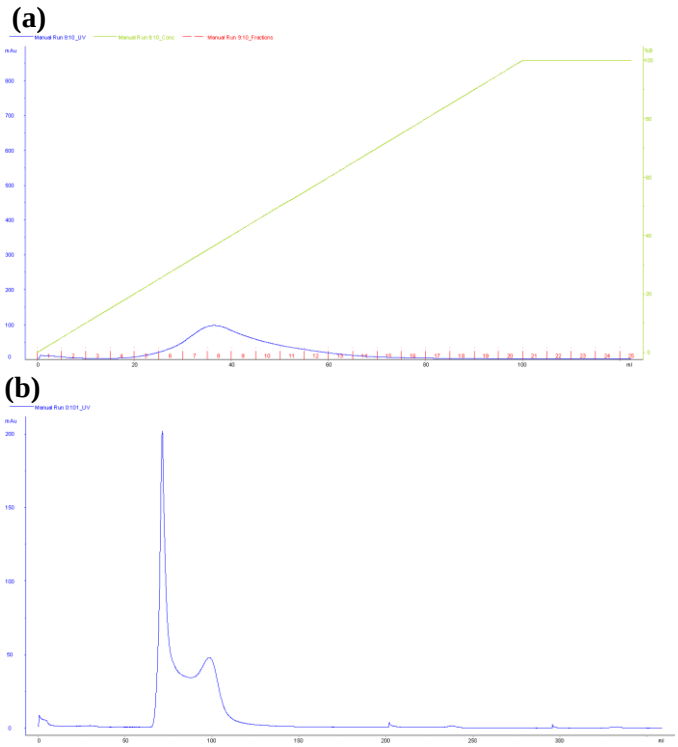


(b) Level of PL_3 expression after 16 hour induction at 16 ℃.
Lane 1: low molecular weight marker (Sigma);
Lane 2: cell culture after 16-hour induction.

Figure 3.6. 15% SDS-PAGE gel of samples taken during the expression of $wtPL_3$ showing the levels of expression after (a) 3 hour induction at 37 ℃, and (b) 16 hour induction at 16 ℃. Proteins were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250.

The optimised expression condition was followed for all subsequent expression trials, and the same purification protocol described in Section 3.2.2 was followed. Surprisingly, very low level of $wtPL_3$ could be eluted from the column. Instead, a large amount of protein was eluted during the column wash

step with 1 M NaOH (Figure 3.7). Since ion exchange chromatography depended the net surface charge of the protein which could be different if the protein was not correctly folded, the next step was therefore try to optimise the refolding condition.



(a) **Chromatogram of the elution step. The peak has a maximum UV absorbance of 80 mAU;**

(b) **Chromatogram of the NaOH washing step. The peak has a maximum UV absorbance of >200 mAU.**

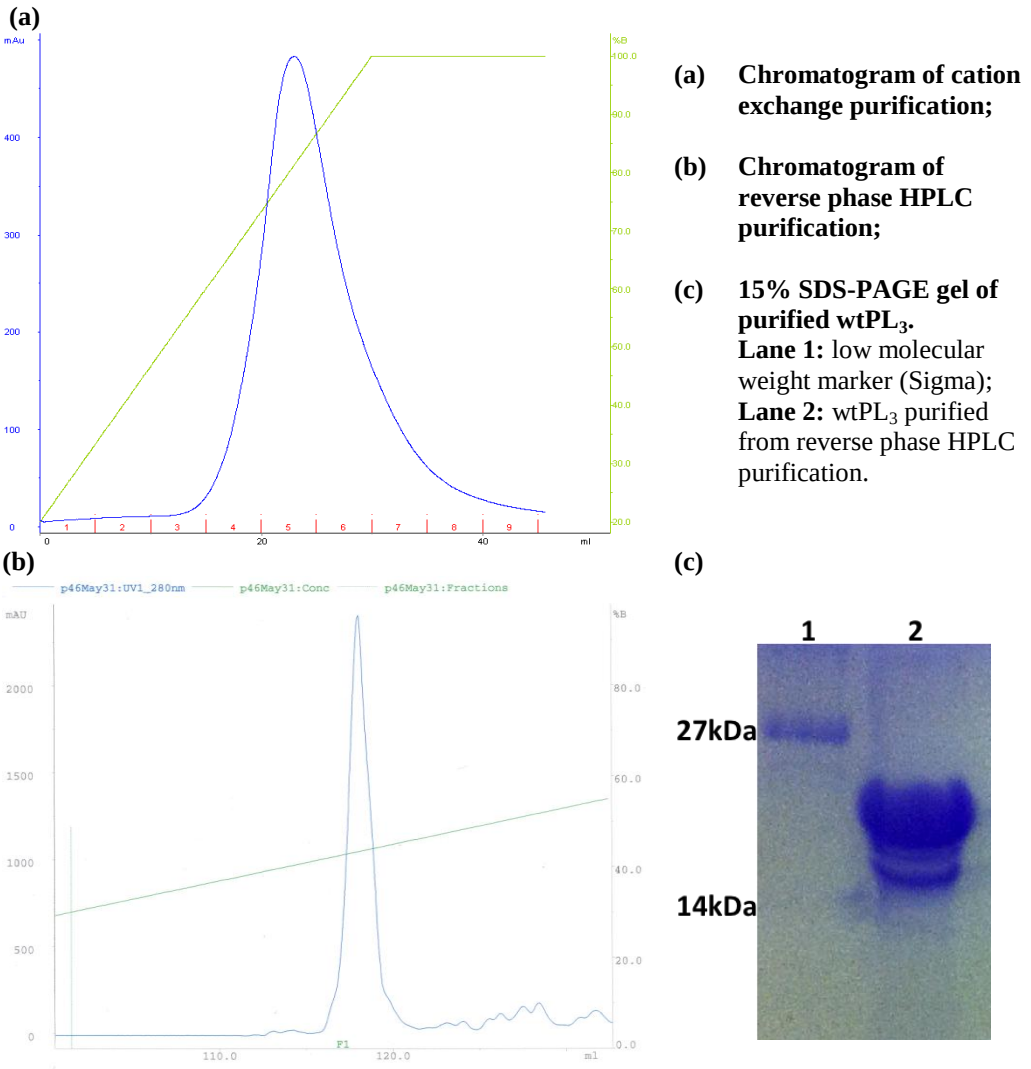
Figure 3.7. Chromatograms of (a) the elution step, and (b) the NaOH washing step of the initial cation exchange purification of wtPL$_3$. Purification only produced a small amount of correctly folded PL$_3$; a large amount of incorrectly folded PL$_3$ was eluted during the NaOH column washing step.

3.3.3.2 Optimisation of the purification condition for pET46-PL$_3$

The protocol as described in Section 3.2.2 was followed for the preparation and solubilisation of wtPL$_3$ inclusion bodies.

The volume of the refolding buffer (0.5 M L-arginine, 0.6 mM oxidized glutathione) was reduced from 1 L to 500 ml to increase the concentration of wtPL$_3$, which may assist correct protein refolding. Subsequent cation exchange purification protocol was as described in Section 3.2.2.

The yield of $wtPL_3$ protein from this purification was significantly improved compared with the original protocol; however, the elution of $wtPL_3$ only began when the concentration of NaCl was close to 0.8 M. To further optimise the purification protocol, the maximum NaCl concentration was increased to 1.5 M during the elution step, whilst the optimised solubilisation and refolding steps, as well as the subsequent HPLC purification protocol were followed. Pure $wtPL_3$ was obtained at the end of the two-step purification (Figure 3.8), approximately 20 mg of $wtPL_3$ per 1 L of bacterial culture.



(a) Chromatogram of cation exchange purification;

(b) Chromatogram of reverse phase HPLC purification;

(c) 15% SDS-PAGE gel of purified $wtPL_3$.
Lane 1: low molecular weight marker (Sigma);
Lane 2: $wtPL_3$ purified from reverse phase HPLC purification.

Figure 3.8. Results of the purifications of $wtPL_3$ from the $BL21C^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 strain. Protein sample was electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. Both (a) the cation exchange purification and (b) the reverse phase HPLC purification showed good separation of the target PL_3 protein. Significant quantities of pure $wtPL_3$ were obtained at the end of the two-step purification, as shown by (c) the 15% SDS-PAGE gel of purified $wtPL_3$.

3.3.4 Mass spectrometry analysis of tagged wild type pET46-PL_3

A single species with a molecular weight of 15539.53 Da was detected by mass spectrometry (Figure 3.9). This value matches the theoretical molecular weight (15539.4 Da) of the tagged PL_3 with the methionine from the tag being removed.

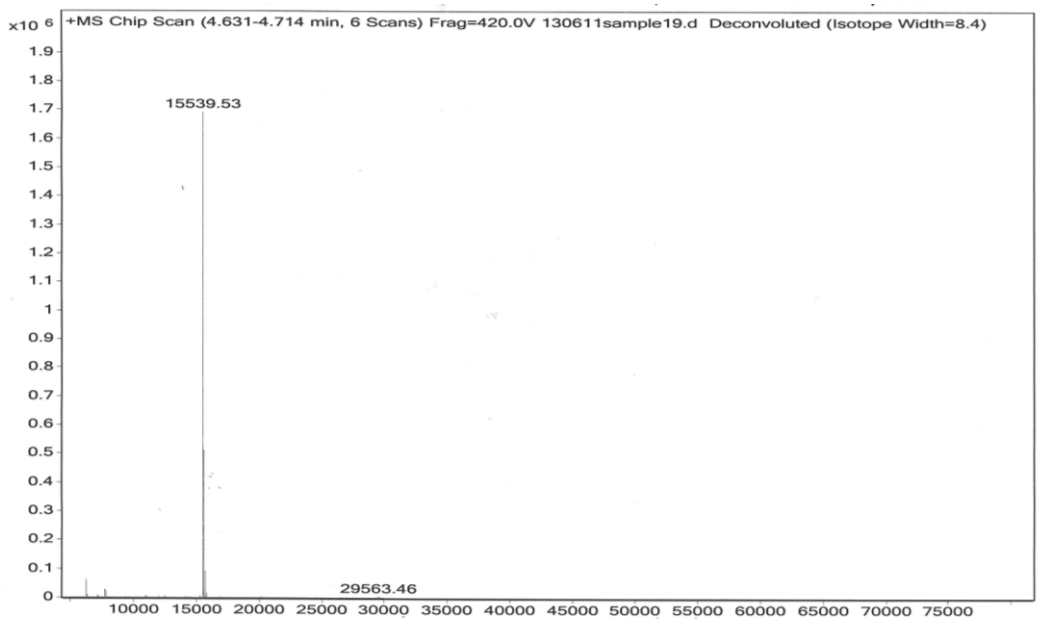


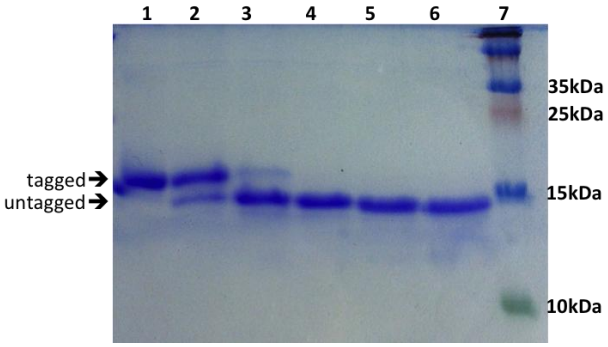
Figure 3.9. Result of the mass spectrometry analysis of $wtPL_3$ sample. A single species with a molecular weight of 15539.53 Da was detected, matching the expected molecular weight of tagged $wtPL_3$.

3.3.5 His-tag removal of $wtPL_3$

Prior to proceeding with the following experiments, the tagged $wtPL_3$ was used in the formation of PL_3·pRI complex where precipitation was observed (Section 4.3.2). Considering the N-terminal hexahistidine tag of PL_3 could hinder the formation of PL_3·pRI complex, experiments were carried out in order to remove the N-terminal hexahistidine tag by enterokinase. Two enterokinase enzymes were used: EKMax from Invitrogen and rEk from NEB.

3.3.5.1 His-tag cleavage by EKMax

A pilot study was carried out to determine the appropriate concentration of EKMax needed for tag cleavage. SDS-PAGE analysis of the pilot study of His-tag cleavage by EKMax is shown in Figure 3.10. All concentrations of EKMax used showed a certain level of digestion; and 0.1 unit of EKMax was the minimal amount tested that was sufficient to completely digest 20 µg of PL_3 protein in 16 hours.

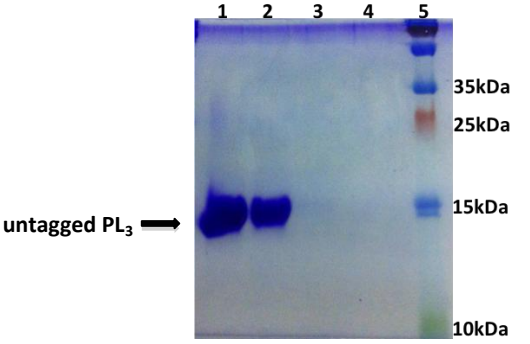


Lane 1: tagged PL_3 with 0 unit of EKMax;
Lane 2: tagged PL_3 with 0.001 unit of EKMax;
Lane 3: tagged PL_3 with 0.01 unit of EKMax;
Lane 4: tagged PL_3 with 0.1 unit of EKMax;
Lane 5: tagged PL_3 with 1 unit of EKMax;
Lane 6: tagged PL_3 with 4 unit of EKMax;
Lane 7: PageRuler Plus Prestained Protein Ladder (Thermo Scientific).

Figure 3.10. 15% SDS-PAGE gel of samples of the pilot study of His-tag cleavage by EKMax. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. 0.1 unit of EKMax was the minimal amount tested that was sufficient to completely digest 20 µg of PL_3 protein in 16 hours.

3.3.5.2 Purification of EKMax digested wtPL_3 by affinity chromatography

The reaction of the His-tag cleavage of wtPL_3 with 0.1 unit of EKMax was scaled up in a linear fashion to produce more untagged wtPL_3. A total amount of 9.2 mg of wtPL_3 was digested with 46 unit of EKMax for 16 hours at room temperature. The EKMax was removed from the PL_3 sample by affinity chromatography with a 1 ml HiTrap Benzamidine FF (High Sub) column from GE Healthcare. Flow through from all steps were collected and analysed by SDS-PAGE (Figure 3.11). Untagged PL_3 did not show sign of interaction with benzamidine, and was present in the fractions of loading and the first washing steps.

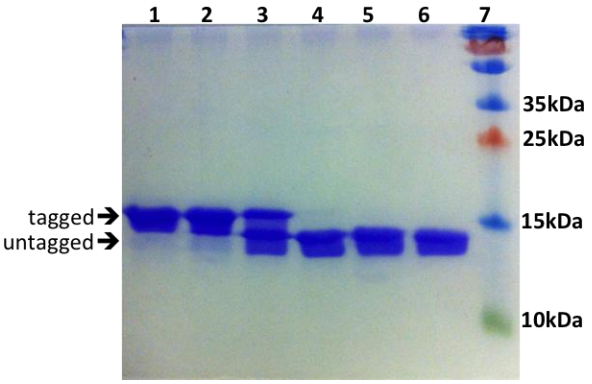


Lane 1: flow through of the loading step;
Lane 2: flow through of the first washing step;
Lane 3: flow through of the second washing step;
Lane 4: eluate;
Lane 5: PageRuler Plus Prestained Protein Ladder (Thermo Scientific).

Figure 3.11. 15% SDS-PAGE gel of samples collected during the benzamidine affinity chromatography purification of EKMax digested PL_3. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. Untagged PL_3 was present in the fractions of loading and the first washing steps. The calculated molecular weight of EkMax is 26.3 kDa; however, it appears on SDS-PAGE gels as 46 kDa due to glycosylation (LaVallie *et al.*, 1993).

3.3.5.3 His-tag cleavage by rEK

Due to a short-term discontinuation of EKMax, rEK from NEB was used instead in future experiments. Another pilot study was carried out using rEK to ensure a similar efficiency as EKMax could be achieved by rEK. SDS-PAGE analysis of the pilot study of His-tag cleavage by rEK is shown in Figure 3.12. All concentrations of rEK used showed a certain level of digestion. 0.1 unit of rEK could cleave the majority of the 20 μg PL_3, indicated by the very weak band corresponding to the tagged PL_3 and a lower strong band of untagged PL_3 protein. The results suggested that 1 unit of rEK was the minimal amount tested that was sufficient to completely digest 20 μg of PL_3 protein in 16 hours.



Lane 1: tagged PL_3 with 0 unit of rEK;
Lane 2: tagged PL_3 with 10^{-3} unit of rEK;
Lane 3: tagged PL_3 with 10^{-2} unit of rEK;
Lane 4: tagged PL_3 with 10^{-1} unit of rEK;
Lane 5: tagged PL_3 with 1 unit of rEK;
Lane 6: tagged PL_3 with 4 unit of rEK;
Lane 7: PageRuler Plus Prestained Protein Ladder (Thermo Scientific).

Figure 3.12. 15% SDS-PAGE gel of samples of the pilot study of His-tag cleavage of wtPL_3 by rEK. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. 1 unit of rEK was the minimal amount tested that was sufficient to completely digest 20 μg of PL_3 protein in 16 hours.

3.3.5.4 Purification of rEK digested wtPL$_3$ by affinity chromatography

The reaction of the His-tag cleavage of wtPL$_3$ with 1 unit of rEK was scaled up in a linear fashion to produce more untagged wtPL$_3$. A total amount of 7.75 mg of wtPL$_3$ was digested with 387.5 unit of rEK for 16 hours at room temperature. The rEk was removed from the PL$_3$ sample by benzamidine affinity chromatography. Flow through from all steps was collected and analysed by SDS-PAGE (Figure 3.13). Untagged PL$_3$ was mainly present in the fractions of loading and the first washing steps.

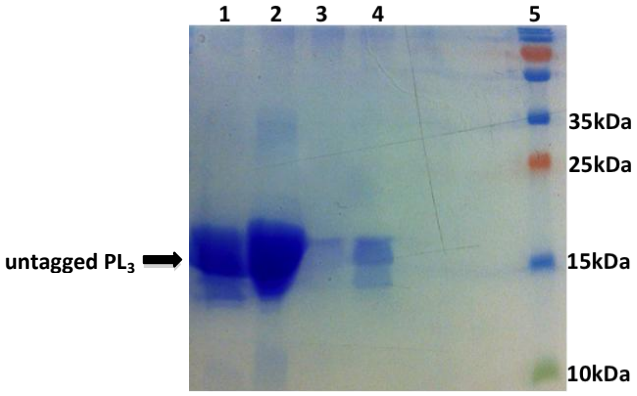


Figure 3.13. 15% SDS-PAGE gel of samples collected during the benzamidine affinity chromatography purification of rEK digested PL$_3$. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. Untagged PL$_3$ was present mainly in the fractions of loading and the first washing steps. The calculated molecular weight of rEk is 26.3 kDa; however, it appears on SDS-PAGE gels as 31 kDa (NEB P8070S).

3.3.5.5 Purification of rEK digested wtPL$_3$ by reverse phase HPLC chromatography

Fractions of the first two steps of the benzamidine affinity chromatography purification were pooled and further purified by reverse phase HPLC chromatography. Protein elution started with ~50% acetonitrile TFA (Figure 3.14).

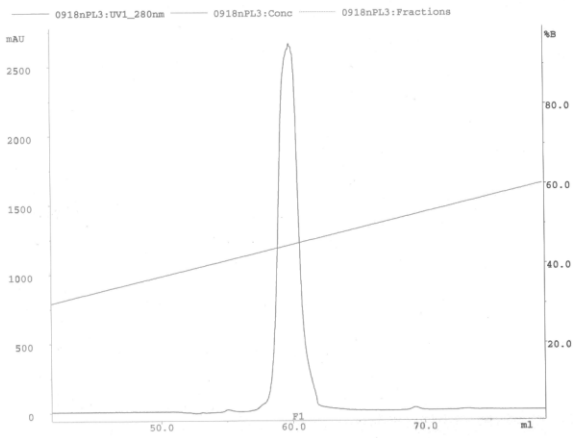


Figure 3.14. Chromatogram of reverse phase HPLC purification of untagged PL_3 showing good separation of the untagged PL_3 protein. The peak has a maximum UV absorbance of ~2700 mAU.

3.3.6 Mass spectrometry analysis of untagged $wtPL_3$

The untagged PL_3 protein sample (10 mg/ml) was analysed by mass spectrometry. The abundant species detected had a molecular weight of 13957.95 Da, which was consistent with that of the untagged (Met-0) PL_3. Another species weighed 15539.53 Da could be detected weakly, which corresponds to that of the tagged PL_3 (Figure 3.15).

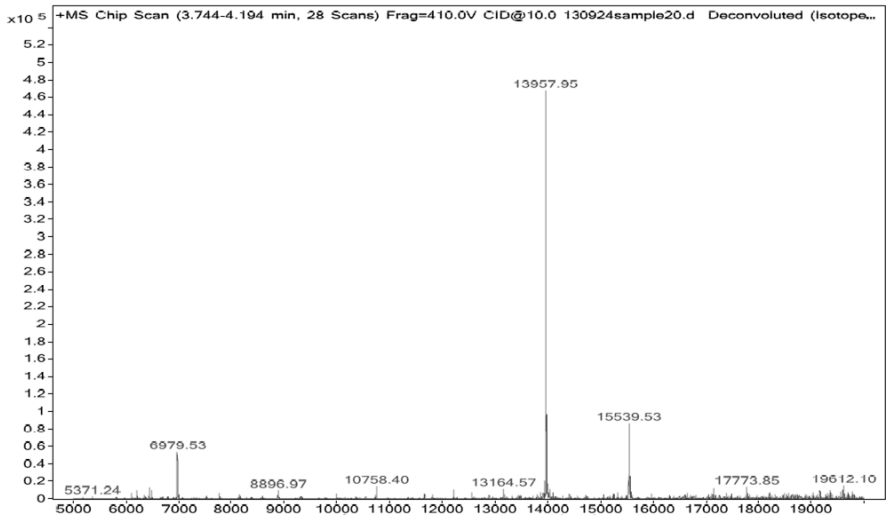


Figure 3.15. Result of the mass spectrometry analysis of untagged $wtPL_3$ sample. The abundant species detected had a molecular weight of 13957.95 Da, consistent with the expected molecular weight of untagged (Met-0) PL_3. Another species weighed 15539.53 Da could be detected weakly, which corresponds to the molecular weight of tagged PL_3.

3.3.7 Construction of PL_3 D80A mutant

3.3.7.1 First attempt

The protocol described in Section 3.2.9 was followed to prepare the pET-46 Ek/LIC-PL_3 D80A construct using the gene of wtPL_3. However, the DNA sequencing result of the ET-46 Ek/LIC-PL_3 D80A plasmid suggested that the constructed mutant was D80M rather than the desired D80A mutant.

3.3.7.2 Second attempt

In the second attempt, two parameters of the PCR recipe were modified: betaine was removed from the recipe to prevent it from over-reducing primer binding to the template; *PfuUltra* II Fusion HS DNA Polymerase was used (Table 3.11). Unfortunately, the mutation from the second attempt was still D80M; therefore it was likely that the primers had the wrong sequence.

Table 3.11. Modified PCR reaction recipes for the amplification of PL_3 D80A gene.

SET 1	Forward primer (10 µM)	1.5 µl
	Reverse primer (10 µM)	1.5 µl
	Template PL_3 DNA	2 µl
	dNTP mix (2 mM each)	5 µl
	25 mM $MgSO_4$	2 µl
	10× PCR Buffer for KOD Hot Start DNA Polymerase	5 µl
	KOD Hot Start DNA Polymerase	1 µl
	H_2O	32 µl
	TOTAL	**50 µl**

SET 2	Forward primer (10 µM)	1 µl
	Reverse primer (10 µM)	1 µl
	Template PL_3 DNA	1 µl
	dNTP mix (2 mM each)	5 µl
	10× *PfuUltra* II Reaction Buffer	5 µl
	PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies)	1 µl
	H_2O	36 µl
	Forward primer (10 µM)	1 µl
	TOTAL	**50 µl**

3.3.7.3 Third attempt

A fresh set of primers (same sequences as shown in Table 3.7) was ordered from Eurofins MWG Operon, and the protocol described in Section 3.2.9 was followed. The plasmid DNA was sequenced, and the result was aligned with the native PL_3 DNA obtained from the NCBI database (Gene ID: 396976) (Figure 3.16). The alignment confirmed that the pET-46 Ek/LIC-PL_3 D80A plasmid was successfully constructed and transformed into Top 10 cells. The plasmid was then used to prepare the expression strain.

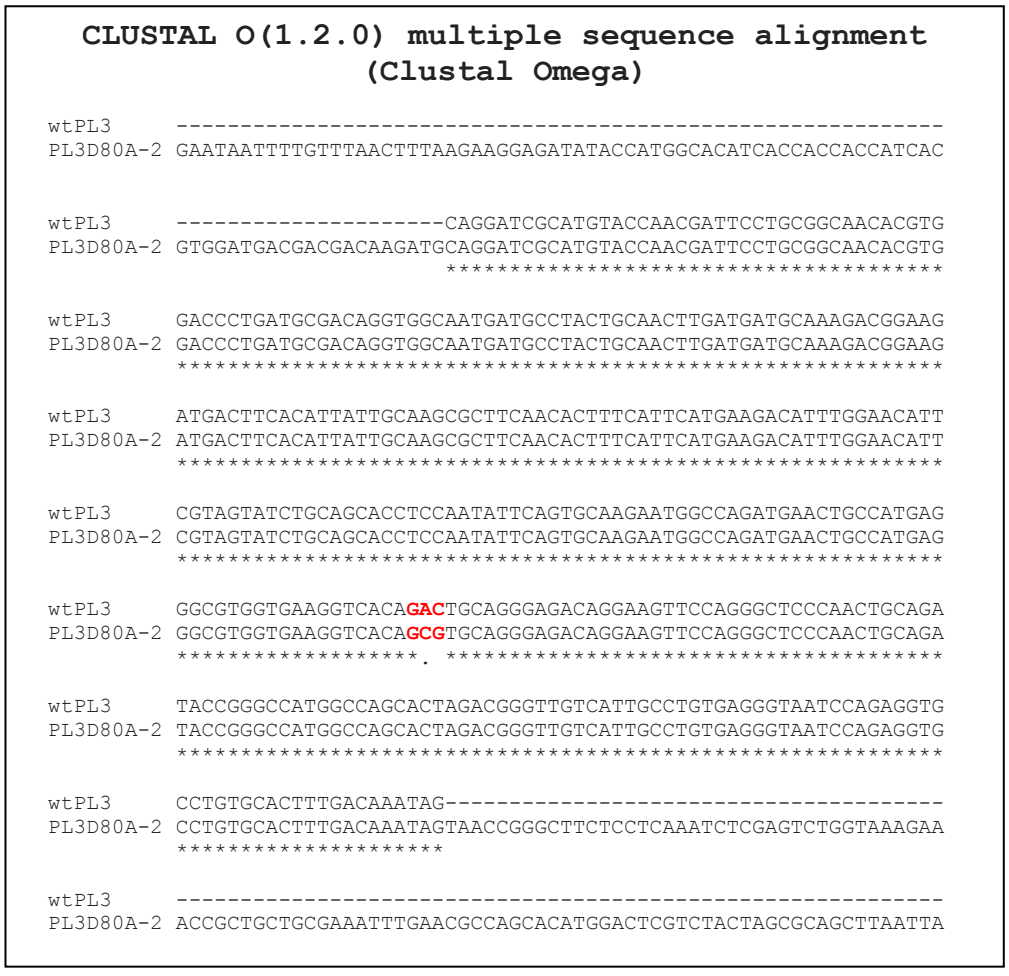
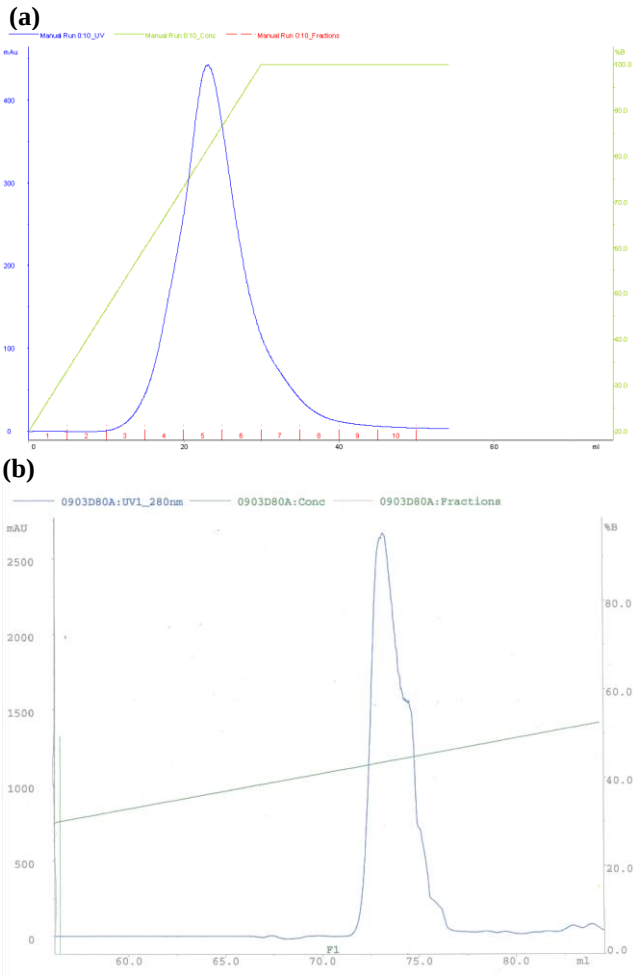


Figure 3.16. Sequence alignment of PL_3 gene from the cloning strain of pET-46 Ek/LIC-PL_3 and the PL_3 gene sequence obtained from NCBI database (Gene ID: 396976). Alignment was performed using *Clustal Omega* online server (Sievers *et al.*, 2011; McWilliam *et al.*, 2013). Bases coloured in red indicate the mutation sites. GAC in the sequence of wild type PL_3 codes for aspartic acid (Asp or D); GCG in the sequence of PL_3 D80A codes for alanine (Ala or A).

3.3.8 Expression and purification of tagged PL_3 D80A mutant

PL_3 D80A was expressed and purified from inclusion bodies following the optimised protocol outlined in Appendix II (Figure 3.17).



(a) Chromatogram of cation exchange purification of PL_3 D80A mutant. The peak has a maximum UV absorbance of ~470mAU.

(b) Chromatogram of reverse phase HPLC purification of PL_3 D80A mutant. The peak has a maximum UV absorbance of ~2700 mAU.

Figure 3.17. Chromatograms of (a) cation exchange purification, and (b) reverse phase HPLC purification of PL_3 D80A mutant, showing good separation of PL_3 D80A protein.

3.3.9 Mass spectrometry analysis of tagged PL_3 D80A mutant

The tagged PL_3 D80A protein sample (8.9 mg/ml) was analysed by mass spectrometry. A single species with a molecular weight of 15495.58 Da was detected (Figure 3.18). This value matches the theoretical molecular weight (15495.4 Da) of the tagged PL_3 D80A mutant with the methionine from the tag being removed.

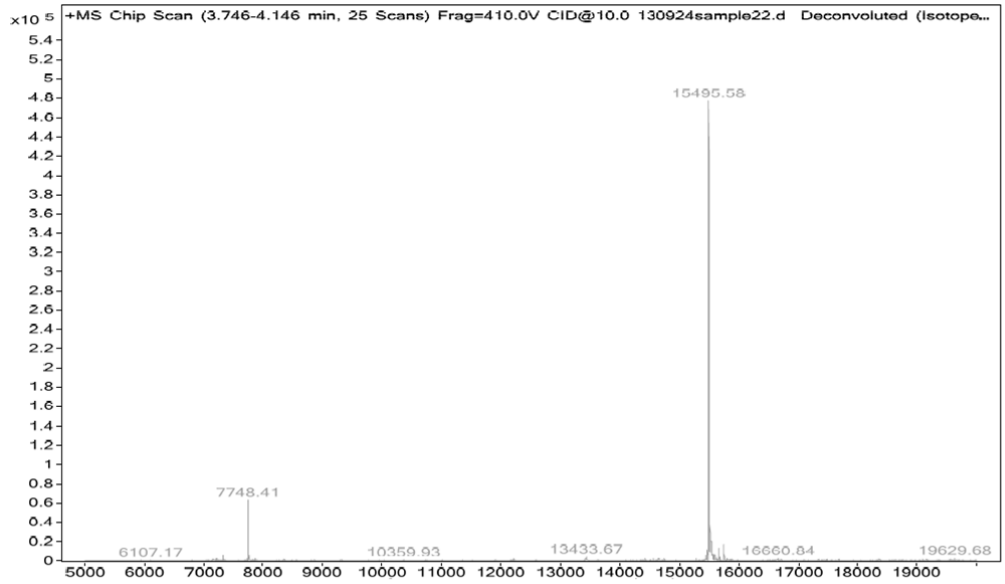
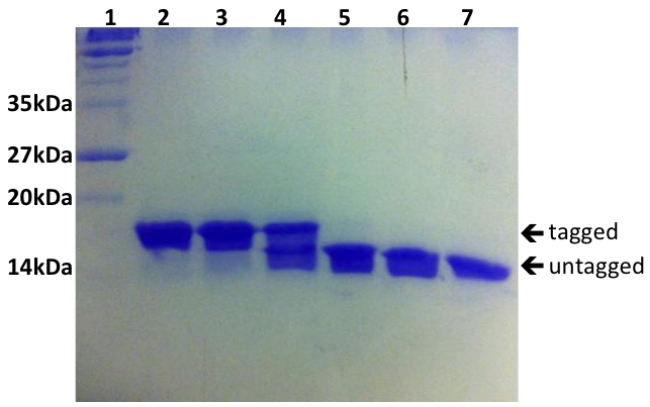


Figure 3.18. Result of the mass spectrometry analysis of tagged PL_3 D80A mutant sample. A single species with a molecular weight of 15495.58 Da was detected, corresponding to the calculated molecular weight of tagged PL_3 D80A.

3.3.10 His-tag removal of PL_3 D80A mutant

3.3.10.1 His-tag cleavage by rEK

Following the protocol outlined in Section 3.2.6, rEK was used to cleave the His tag from purified PL_3 D80A. SDS PAGE analysis confirmed that cleavage was successful, as shown in Figure 3.19. All concentrations of rEK used showed a certain level of digestion. Similar to the results of His-tag cleavage of wtPL_3 by rEK, 0.1 unit of rEK could cleave the majority of the 20 μg PL_3 D80A protein, and 1 unit of rEK was the minimal amount tested that was sufficient to completely digest 20 μg of PL_3 D80A protein in 16 hours.

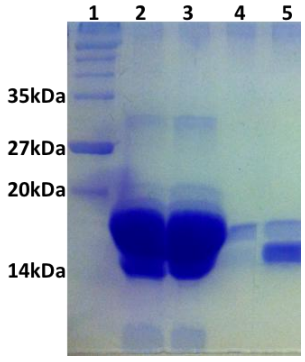


Lane 1: low molecular weight marker (Sigma);
Lane 2: tagged PL_3 D80A with 0 unit of rEK;
Lane 3: tagged PL_3 D80A with 10^{-3} unit of rEK;
Lane 4: tagged PL_3 D80A with 10^{-2} unit of rEK;
Lane 5: tagged PL_3 D80A with 10^{-1} unit of rEK;
Lane 6: tagged PL_3 D80A with 1 unit of rEK;
Lane 7: tagged PL_3 D80A with 4 unit of rEK.

Figure 3.19. 15% SDS-PAGE gel of samples of the pilot study of His-tag cleavage of PL_3 D80A mutant by rEK. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. 1 unit of rEK was the minimal amount tested that was sufficient to completely digest 20 µg of PL_3 D80A protein in 16 hours.

3.3.10.2 Purification of rEK digested PL_3 D80A mutant by affinity chromatography

A total amount of 10.12 mg of concentrated PL_3 D80A sample was digested by rEk, which was subsequently removed from the protein sample by benzamidine affinity chromatography. Flow through from all steps were analysed by SDS-PAGE (Figure 3.20). Untagged PL_3 D80A protein did not show sign of interaction with benzamidine, and was present mainly in the fractions of loading and the first washing steps.



Lane 1: low molecular weight marker (Sigma);
Lane 2: flow through of the loading step;
Lane 3: flow through of the first washing step;
Lane 4: flow through of the second washing step;
Lane 5: eluate.

Figure 3.20. 15% SDS-PAGE gel of samples collected during the benzamidine affinity chromatography purification of rEK digested PL_3 D80A mutant. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. Untagged PL_3 D80A was present mainly in the flow through of loading and the first washing steps. The calculated molecular weight of rEk is 26.3 kDa; however, it appears on SDS-PAGE gels as 31 kDa (NEB P8070S).

3.3.10.3 Purification of rEK digested PL_3 D80A mutant by reverse phase HPLC chromatography

Fractions collected from the first two steps of the benzamidine affinity chromatography purification were pooled and further purified by reverse phase HPLC chromatography. Untagged PL_3 D80A protein elution started with ~50% acetonitrile :TFA (Figure 3.21).

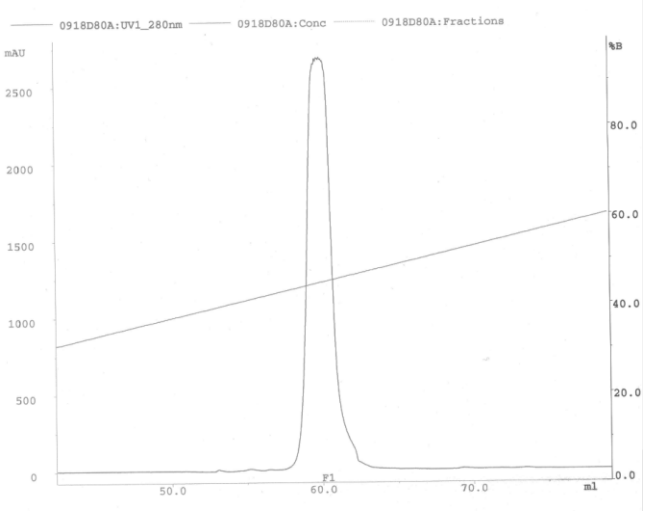


Figure 3.21. Chromatogram of reverse phase HPLC purification of untagged PL_3 D80A mutant showing good separation of the untagged PL_3 D80A protein. The peak has a maximum UV absorbance of ~2700 mAU.

3.3.11 Mass spectrometry analysis of untagged PL_3 D80A mutant

The untagged PL_3 D80A protein sample (10.5 mg/ml) was analysed by mass spectrometry. A single species with a molecular weight of 13913.94 Da was detected (Figure 3.22), which was consistent with the theoretical molecular weight (13913.8 Da) of the untagged (Met-0) PL_3 D80A mutant.

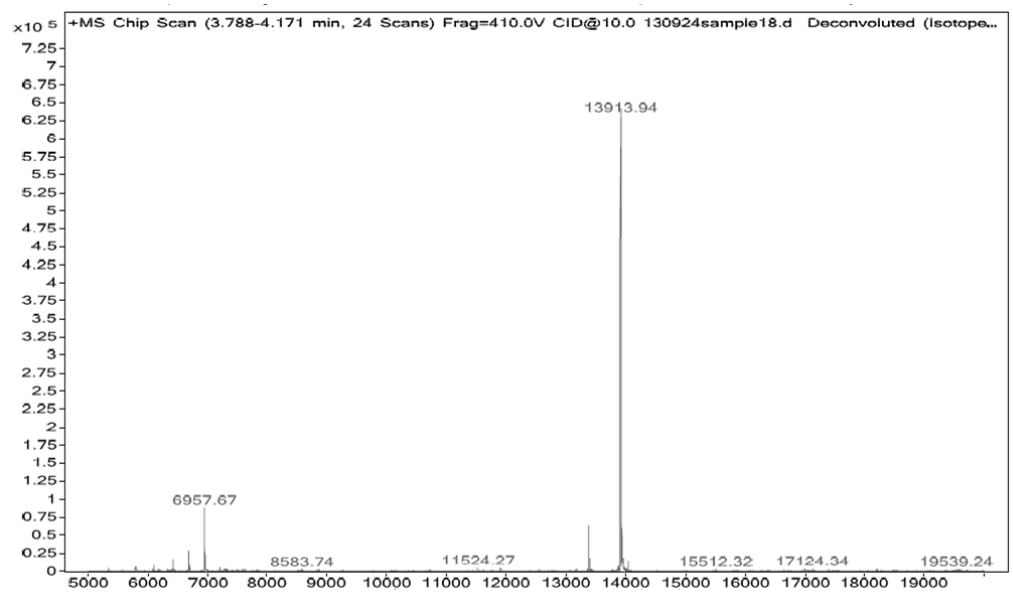


Figure 3.22. Result of the mass spectrometry analysis of untagged PL_3 D80A mutant sample. A single species with a molecular weight of 13913.94 Da was detected, matching the calculated molecular weight of the untagged (Met-0) PL_3 D80A mutant.

3.3.12 Ribonucleolytic activity assay of wtPL_3 and PL_3 D80A mutant

The ribonucleolytic activities of tagged and untagged wtPL_3 and PL_3 D80A protein samples were analysed by conducting the tRNA cleavage assays as described by Shapiro *et al.* (1987). This assay was originally performed to measure the ribonucleolytic activity of hANG, and has been widely used as a 'robust method' to assess the activities of other RNase A homologues, especially those purified from inclusion bodies. Since ANG possesses ribonucleolytic activity, it was used as a control to assess the relative ribonucleolytic activity of PL_3 and PL_3 D80A mutant proteins.

The result of the initial RNase assay experiment revealed that native PL_3 protein is more active than native hANG with the same concentration. The four PL_3 samples used in the subsequent assay experiments were therefore diluted to 0.04 μg/ml whilst the concentration of hANG was kept as 40 μg/ml, which reduced the molar concentrations of the four PL_3 samples by 1,000-fold.

The absorbance readings of all protein samples were plotted on the same graph (Figure 3.23). The relative activity of each test sample was calculated using the method described by Crabtree *et al.* (2007).

The results (Figure 3.23) showed that all four PL_3 samples were active towards tRNA; and both tagged and untagged $wtPL_3$ were more active towards tRNA than the two PL_3 D80A protein samples. The relative activities of test proteins suggested that the D80A substitution slightly reduced the ribonucleolytic activity of PL_3. However, since the nucleic acid composition of the yeast tRNA was not specified, more analyses need to be performed in order to draw a solid conclusion on the effect of the D80A mutation on the general ribonucleolytic activity of PL_3.

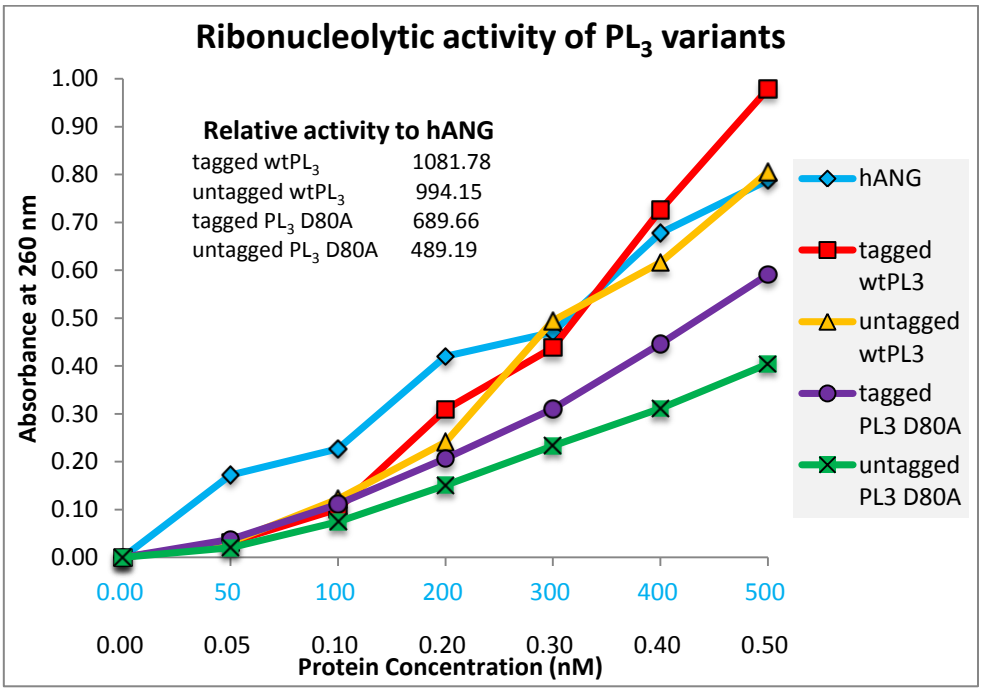


Figure 3.23. Ribonucleolytic activities of PL_3 and PL_3 D80A samples. Y-axis is the absorbance of the supernatant of the reaction sampless at 260 nm; x-axis is the protein concentration at nM, values in blue are the concentrations of hANG, values in black are the concentrations of PL_3 and PL_3 D80A samples. All PL_3 samples showed ribonucleolytic activity towards yeast tRNA. The activities for 0.3 nM to 0.5 nM of each test protein were multiplied 1,000-fold, divided by the activity of hANG, and averaged to give the final relative activity. Wild type PL_3 samples were ~1000× more active than hANG, whilst PL_3 D80A samples were ~500× more active than hANG.

3.3.13 Crystallisation of wtPL_3 and PL_3 D80A mutant

Four protein samples, namely tagged wtPL_3, untagged wtPL_3, tagged PL_3 D80A, and untagged PL_3 D80A were used in the screening for crystallisation conditions.

3.3.13.1 PCTs and initial screening for crystallisation conditions

PCTs were first performed, the results of which suggested that 10 mg/ml was an appropriate starting concentration for all four samples tested. Microcrystals were observed in several conditions from the subsequent 96-well screenings (Table 3.12).

Table 3.12. Initial crystallisation conditions of wtPL$_3$ and PL$_3$ D80A mutant.

tagged wtPL$_3$	
Structure screen I & II – C3	0.1 M sodium HEPES pH 7.5, 1.5 M lithium sulphate
Structure screen I & II – E5	0.1 M Tris pH 8.4, 25% v/v *tert*-butanol
Structure screen I & II – E9	0.1 M Tris pH 8.5, 20% v/v ethanol
Structure screen I & II – H3	0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulphate, 30% w/v PEG 2000 MME
JCSG-plus – C7	0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate dihydrate, 10% w/v PEG 3000
Morpheus – E11	0.1 M Buffer System 3[1] pH 8.5, 0.12 M Ethylene Glycols[2], 30% v/v GOL_P4K[3]
Morpheus – F10	0.1 M Buffer System 3 pH 8.5, 0.12 M Monosaccharides[4], 30% v/v EDO_P8K[5]
Morpheus – F11	0.1 M Buffer System 3 pH 8.5, 0.12 M Monosaccharides, 30% v/v GOL_P4K
untagged wtPL$_3$	
JCSG-plus – E7	0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate dihydrate, 10% v/v 2-Propanol
tagged PL$_3$ D80A mutant	
Structure screen I & II – A3	0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulphate, 25% v/v PEG 4000
Structure screen I & II – A7	0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate, 30% v/v PEG 4000
Structure screen I & II – H3	0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulphate, 30% v/v PEG 2000 MME
Structure screen I & II – H12	1.6 M sodium citrate pH 6.5
JCSG-plus – E1	0.1 M sodium cacodylate pH 6.5, 1.0 M sodium citrate tribasic dihydrate
untagged PL$_3$ D80A mutant	
Structure screen I & II – B7	0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate, 18% PEG 8000
Structure screen I & II – E8	0.1 M Tris pH 8.5, 50% v/v MPD
JCSG-plus – A6	0.1 M phosphate/citrate pH 4.2, 0.2 M lithium sulphate, 20% PEG 1000

[1] Composition of Buffer System 3: Tris (base) and Bicine;

[2] Composition of Ethylene Glycols: di-ethyleneglycol, tri-ethyleneglycol, tetra-ethyleneglycol, and penta-ethyleneglycol;

[3] Composition of GOL_P4K: glycerol and PEG 4000;

[4] Composition of Monosaccharides: D-glucose, D-mannose, D-galactose, L-fucose, D-xylose, and N-Acetyl-D-glucosamine;

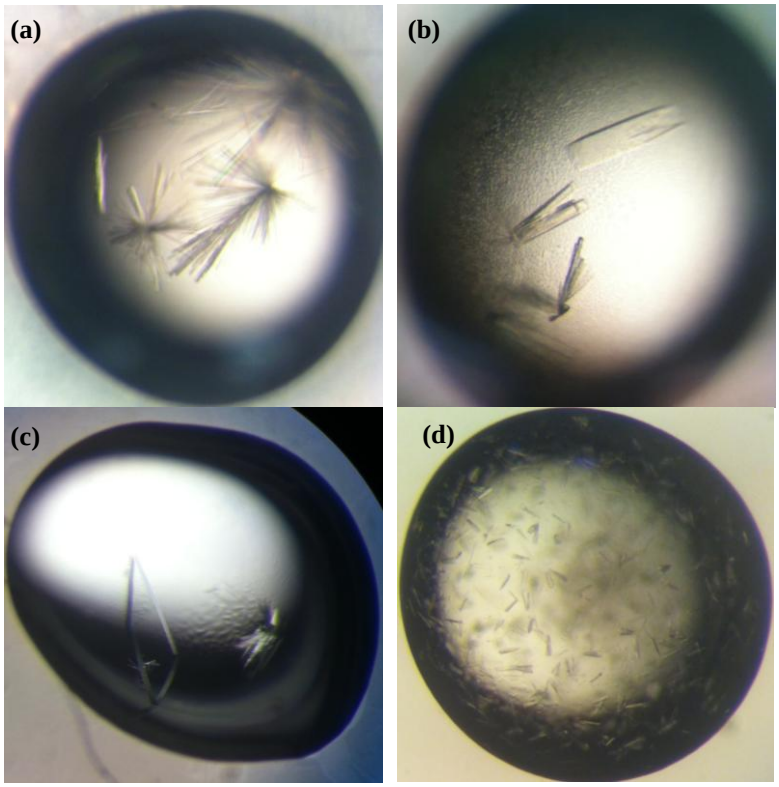
[5] Composition of EDO_P8K: ethylene glycol and PEG 8000.

3.3.13.2 Optimisation of crystallisation conditions for wtPL$_3$ and PL$_3$ D80A

Larger crystals of tagged wtPL$_3$ and tagged PL$_3$ D80A could be obtained from the optimised crystallisation conditions listed in Table 3.13. Examples of these crystals are shown in Figure 3.24.

Table 3.13. Optimised crystallisation conditions of wtPL$_3$ and PL$_3$ D80A.

tagged wtPL$_3$
0.1 M Tris pH 8.5, 20 – 10% v/v ethanol
0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate dihydrate, 15% w/v PEG 3000
0.1 M Morpheus Buffer System 3 pH 8.5, 0.12 M Monosaccharides, 30% – 15 v/v EDO_P8K
tagged PL$_3$ D80A mutant
0.1 M sodium cacodylate pH 6.5, 1.0 M sodium citrate tribasic dihydrate



(a) – **(c)** crystals of tagged wtPL$_3$ protein;

(d) crystals of tagged PL$_3$ D80A protein.

Figure 3.24. Examples of wtPL$_3$ and PL$_3$ D80A protein crystals obtained from optimised conditions.

3.3.14 Crystallisation of $wtPL_3$ dUMP and PL_3 D80A dCMP complexes

3.3.14.1 Soaking tagged $wtPL_3$ crystals with $(UpA)_5$ and dUMP ligands

$(UpA)_5$ oligonucleotide was initially used in the soaking experiment, as UpA was used in the substrate preference experiments of RNase 4. However, preliminary analysis on the datasets suggested no ligand was bound in the crystals.

Considering the fact that $(UpA)_5$ is a hydrolysable ligand of PL_3, the likelihood of successfully co-crystallising PL_3 with this ligand would be low. A group of non-hydrolysable ligands commonly used in the studies of RNases is the 2'-deoxy-2'-fluorouribose substrate analogues, such as 2'-deoxy-2'-fluorouridylyl-3',5'-adenosine (2'-F-dUpA). As these substrates would have to be specially synthesised, a commercially available mononucleotide 2'-deoxyuridine-5'-monophosphate (dUMP) was firstly chosen.

Datasets were collected using dUMP soaked $wtPL_3$ crystals; however, analyses of the datasets suggested that the crystals were ligand-free. As soaking experiments did not produce complex crystals, co-crystallisation was then performed.

3.3.14.2 PCTs and initial screenings for crystallisation conditions

To study the structural basis of RNase 4 substrate recognition, especially the role Asp-80 plays, we chose dUMP and dCMP as the ligands for $wtPL_3$ and PL_3 D80A respectively.

PCTs were first performed using protein samples (tagged $wtPL_3$, untagged $wtPL_3$ tagged PL_3 D80A and untagged PL_3 D80A) mixed with their corresponding ligands in a 1:1.2 molar ratio. The results of PCT suggested that

a final concentration of 9 mg/ml was an appropriate starting point for all four samples tested. Microcrystals were observed in several conditions from the subsequent 96-well screenings (Table 3.14).

Table 3.14. Initial crystallisation conditions of wtPL_3 dUMP and PL_3 D80A dCMP.

tagged wtPL_3 dUMP	
PACT permier – A5	0.1 M SPG buffer pH 8.0, 25% w/v PEG 1500
JCSG-plus – C7	0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate, 10% w/v PEG3000
JCSG-plus – D4	0.1 M sodium acetate pH 4.5, 0.2 M lithium sulphate, 30% w/v PEG8000
JCSG-plus – E7	0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate dihydrate, 10% v/v 2-propanol
untagged wtPL_3 dUMP	
JCSG-plus – E7	0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate dihydrate, 10% v/v 2-propanol
ProPlex – E7	0.1 M sodium phosphate pH 6.5, 12 % PEG 8000
Stura + MacroSol – G7	0.2 M ammonium sulphate, 24% w/v PEG 8000
tagged PL_3 D80A dCMP	
Structure screen I & II – A7	0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate, 30% v/v PEG 4000
Structure screen I & II – A9	0.1 M sodium citrate pH 5.6, 20% v/v 2-propanol, 20% w/v PEG4000
Structure screen I & II – B7	0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate, 18% PEG8000
ProPlex – B5	0.1 M sodium HEPES pH 7.5, 0.1 M magnesium chloride, 10% w/v PEG 4000

Additive screenings

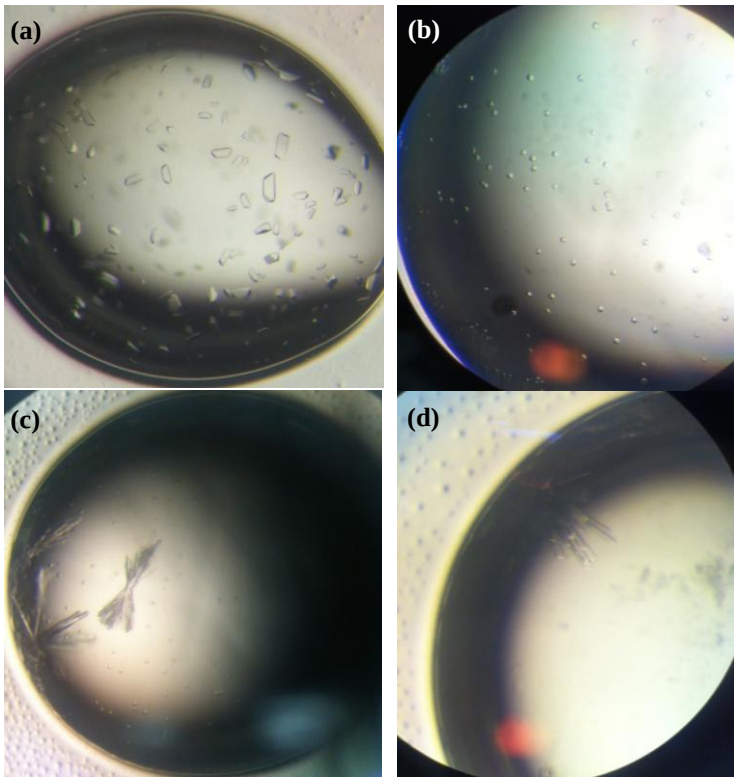
Since the crystals of tagged PL_3 D80A dCMP initially observed in the conditions listed in Table 3.14 were thin needle shaped, which could be different to fish out from the drops and more easily broken by high energy X-rays, therefore additive screenings were also performed for the tagged PL_3 D80A dCMP complex. A few conditions from the additive screenings produced crystals; however they still showed morphology of thin and long needles, suggesting these conditions were not the best choices to optimise.

3.3.14.3 Optimisation of crystallisation conditions for wtPL$_3$ dUMP and PL$_3$ D80A dCMP complexes

Crystals of untagged wtPL$_3$ dUMP and tagged PL$_3$ D80A dCMP complexes could be obtained from the optimised crystallisation conditions listed in Table 3.15. Examples of these crystals are shown in Figure 3.25.

Table 3.15. Optimised crystallisation conditions of wtPL$_3$ dUMP and PL$_3$ D80A dCMP.

untagged wtPL$_3$ dUMP
0.1 M sodium phosphate pH 6.5, 12 – 8% PEG 8000
tagged PL$_3$ D80A dCMP
0.1 M sodium citrate pH 5.6, 20 –15% v/v 2-propanol, 20 – 15% w/v PEG4000
0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate, 26 – 20% v/v PEG 4000



(a) crystals of untagged wtPL$_3$ dUMP;

(b) – **(d)** crystals of tagged PL$_3$ D80A dCMP

Figure 3.25. Examples of wtPL$_3$ dUMP and PL$_3$ D80A dCMP protein crystals obtained from optimised conditions.

3.3.15 Data collection and structure solutions of $wtPL_3$, PL_3 D80A mutant, $wtPL_3$ ·dUMP and PL_3 D80A ·dCMP complexes

3.3.15.1 Data collection and structure solutions of $wtPL_3$ and PL_3 D80A

The crystals of tagged $wtPL_3$ and tagged PL_3 D80A shown in Figure 3.24c and Figure 3.24d diffracted to 1.17 Å (Figure 3.26) and ~2.8 Å (Figure 3.27) respectively.

The structure of $wtPL_3$ was solved in space group I2 (given by *AIMLESS*, equivalent to space group C2) at 1.25 Å with one molecule in the asymmetric unit (ASU); and the structure of PL_3 D80A was solved in space group P1 at 2.9 Å with four molecules in the ASU. Data collection and refinement statistics are shown in Table 3.16; Ramachandran plots of both structures are shown in Figure 3.28 and 3.29.

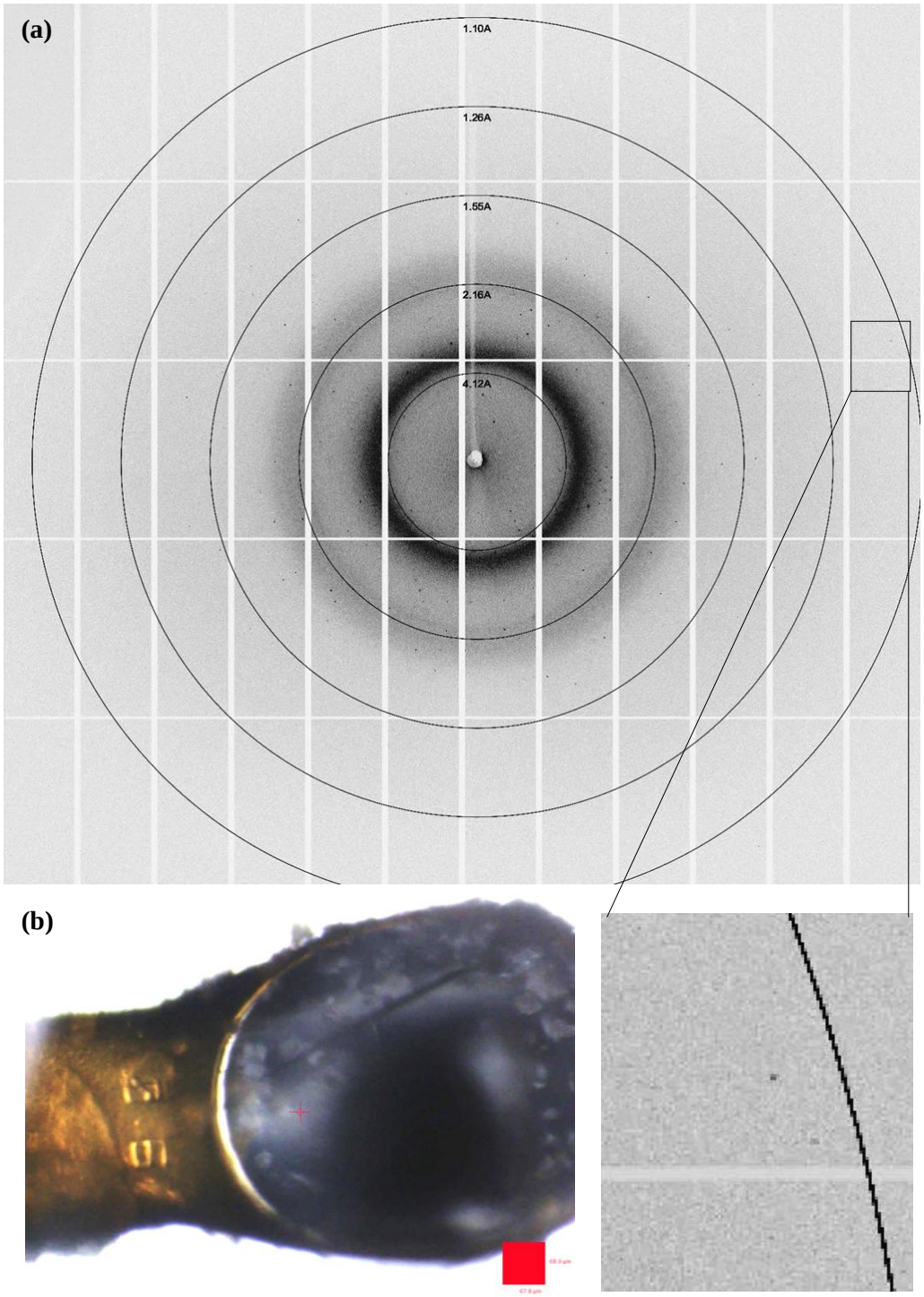


Figure 3.26. Crystals and X-ray diffraction picture of wtPL_3. (a) Diffraction image of a wtPL_3 crystal showing diffractions at 1.17 Å; (b) the PL_3 crystal mounted for data collection, the size of the red square is 67.8µm×68.9µm.

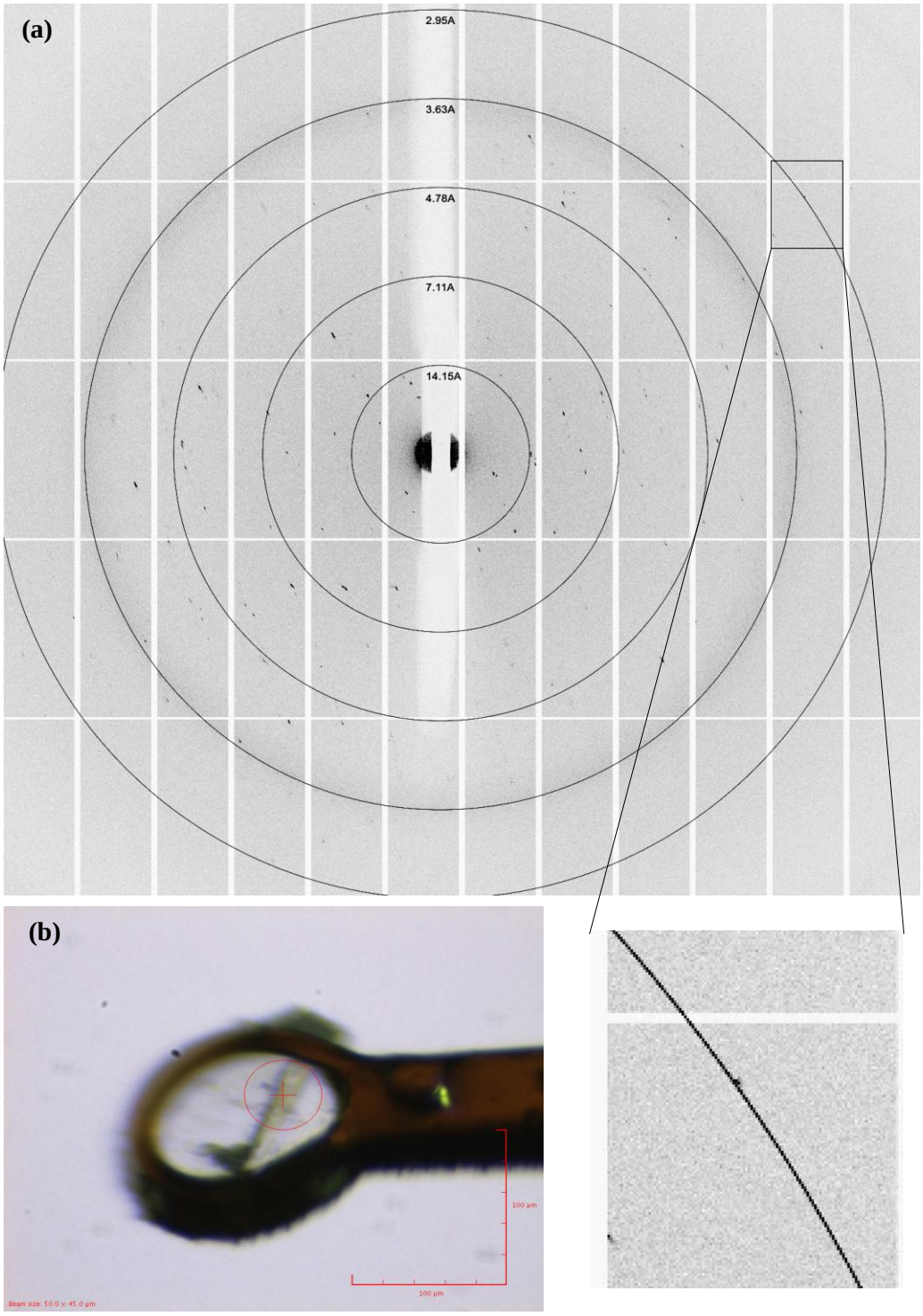


Figure 3.27. Crystals and X-ray diffraction picture of PL_3 D80A mutant. (a) Diffraction image of a PL_3 D80A mutant crystal showing diffractions at 2.9 Å; (b) the PL_3 D80A crystal mounted for data collection, the size of the red square is 100µm×100µm.

Table 3.16. Data collection and refinement statistics of $wtPL_3$ and PL_3 D80A.

	$wtPL_3$	**PL_3 D80A mutant**
A. Data Collection Statistics		
Space Group	I2	P1
Cell	a = 36.29 Å, b = 38.19 Å, c = 86.92 Å α = γ = 90 °; β = 97.68 °	a = 30.83 Å, b = 61.82 Å, c = 68.89 Å α = 92.04 °; β = 101 °; γ = 101.2 °
No. of molecules in the ASU	1	4
Resolution Range (Å)	50 – 1.25	50 – 2.90
R_{sym} [a]	0.059 (0.390)*	0.137 (0.201)*
I/σ(I)	15.9 (4.4)*	8.0 (2.0)*
Completeness (%)	99.9 (100.0)*	73.5 (75.5)*
Redundancy	6.6 (6.6)*	1.6 (1.5)*
B. Refinement Statistics		
Resolution range (Å)	50 – 1.25	50 – 2.90
No. of reflections	215315	12420
Unique No. of reflections	32743	7962
R_{cryst} [b]	0.13	0.21
R_{free} [c]	0.15	0.28
Number of non-H atoms:		
Protein	1017	3853
Ligand	2 ethylene glycol	0
Water	130	0
Average temperature factor (Å^2)	19.5	65.8
RMSD for bond lengths (Å)	0.01	0.009
RMSD for bond angles ( )	1.45	1.28

* Values in parentheses are for the outer shells, which is 1.27 – 1.25Å, and 3.08 – 2.90Å respectively.

[a] $R_{sym} = \frac{\sum_h \sum_i |I_i(h) - I(h)|}{\sum_h \sum_i I_i(h)}$, where $I_i(h)$ and $I(h)$ are the i^{th} and the mean measurements of the intensity of reflection *h*, respectively.

[b] $R_{cryst} = \frac{\sum_h |F_o - F_c|}{\sum_h F_o}$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection *h*, respectively.

[c] R_{free} is equal to R_{cryst} for a randomly selected 5% reflections not used in the refinement.

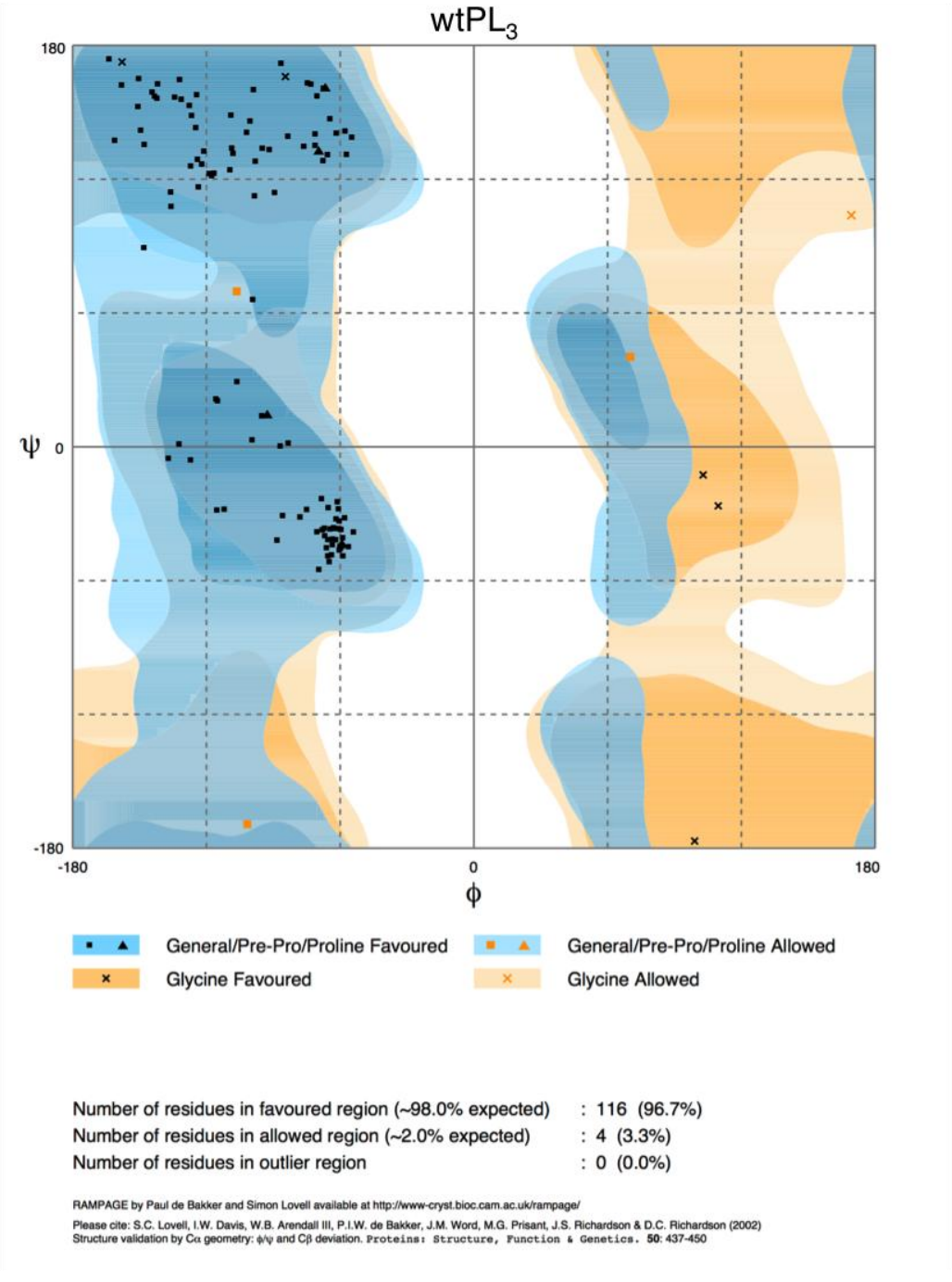


Figure 3.28. Ramachandran plot of wtPL$_3$ structure, which has one molecule in the ASU. 96.7% of the residues are lying in the most favoured regions. The Ramachandran plot was calculated using *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999).

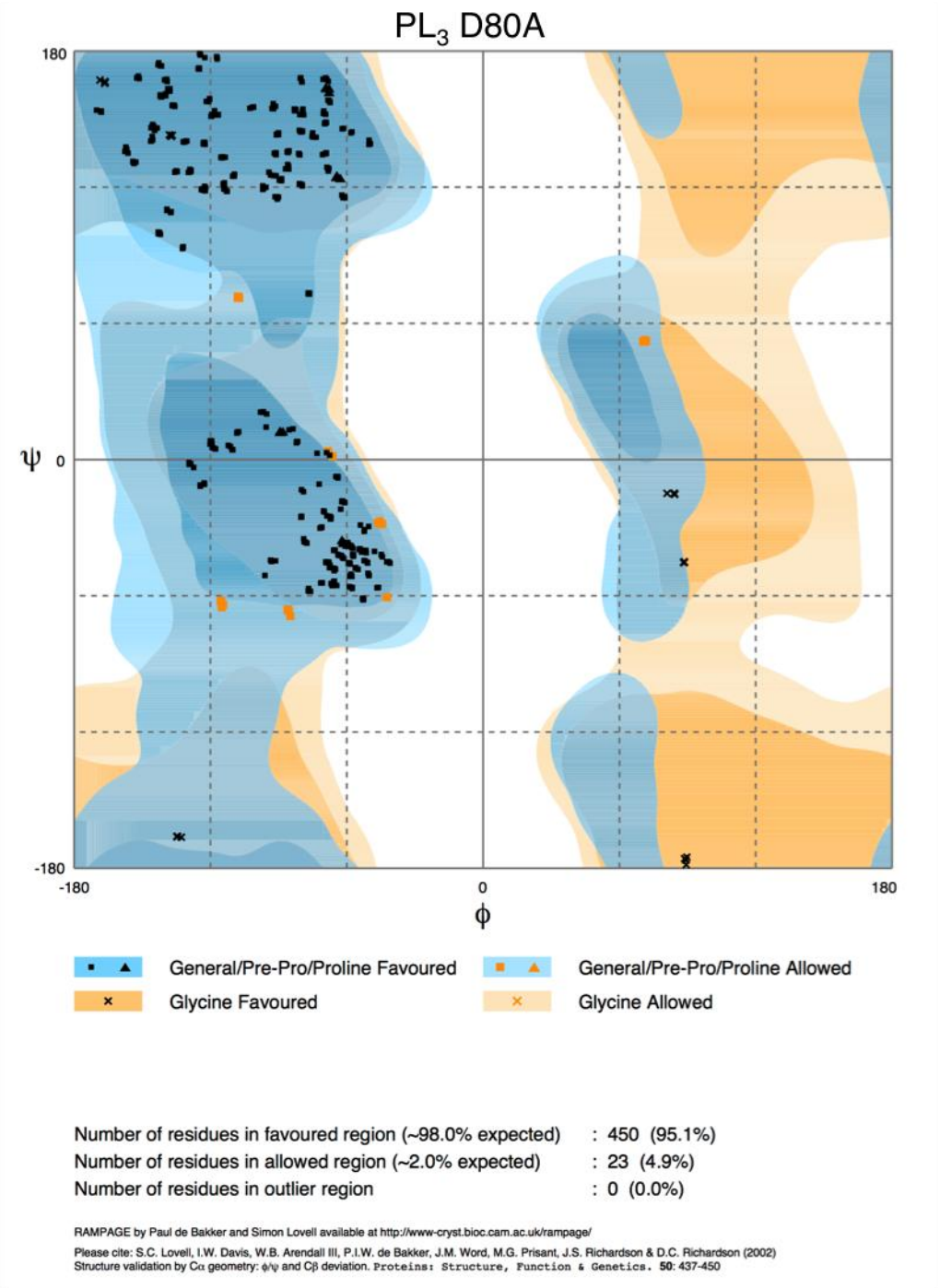


Figure 3.29. Ramachandran plot of PL_3 D80A mutant structure, which has four molecules in the ASU. 94.5% of the residues are lying in the most favoured regions. The Ramachandran plot was calculated using *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999).

3.3.15.2 Data collection and structure solutions of wtPL$_3$ dUMP and PL$_3$ D80A dCMP complexes

The crystals of untagged wtPL$_3$ dUMP and tagged PL$_3$ D80A dCMP shown in Figure 3.25a and Figure 3.25d diffracted to ~2.1 Å (Figure 3.30) and ~2.25 Å (Figure 3.31) respectively.

The structure of wtPL$_3$ dUMP was solved in space group $P2_12_12_1$ at 2.28 Å with four molecules in the ASU; and the structure of PL$_3$ D80A dCMP was solved in space group P1 at 2.6 Å with four molecules in the ASU. Data collection and refinement statistics are shown in Table 3.17; Ramachandran plots of both structures are shown in Figure 3.32 and 3.33.

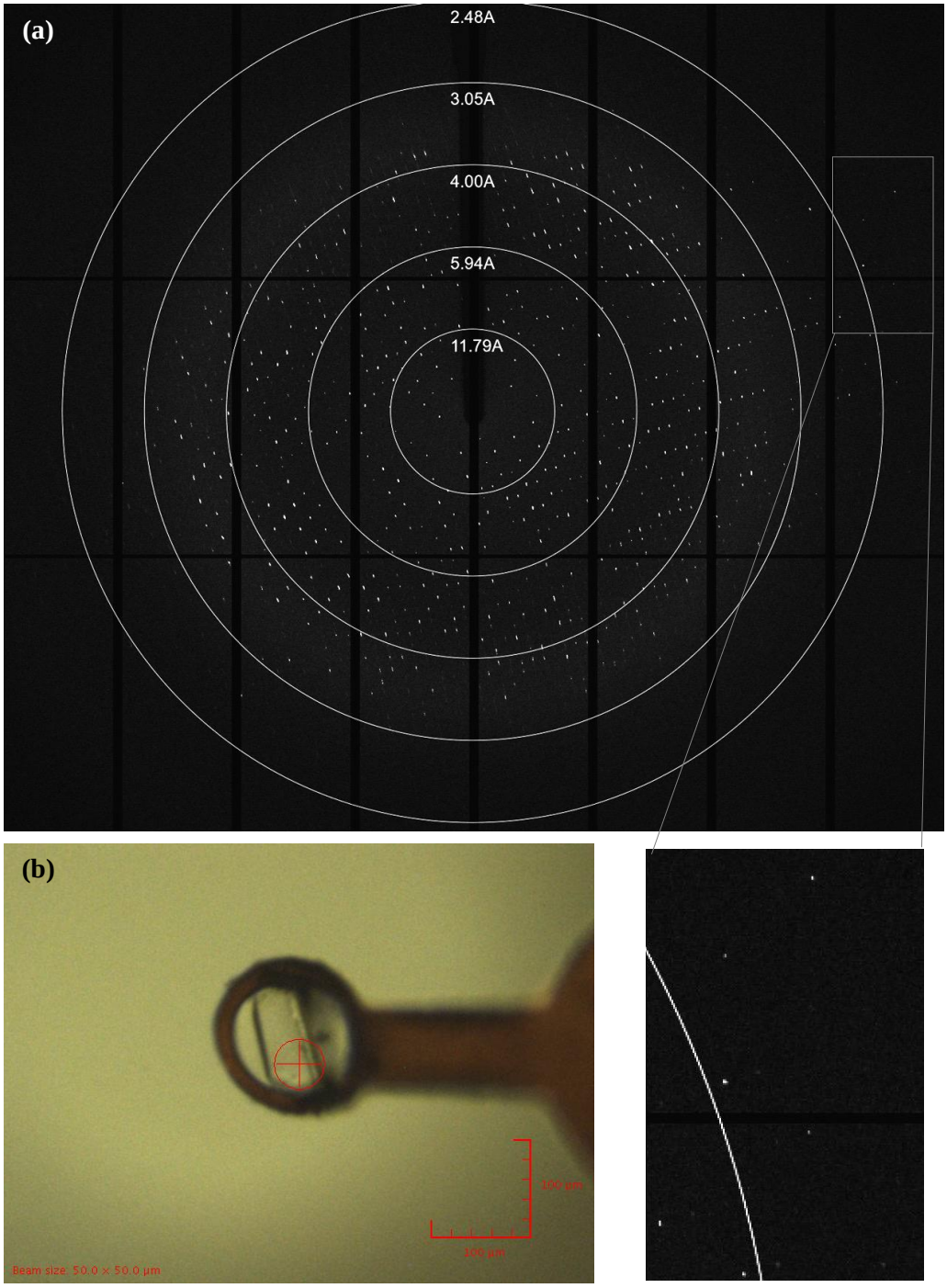


Figure 3.30. Crystals and X-ray diffraction picture of wtPL$_3$·dUMP complex. (a) Diffraction image of a wtPL$_3$·dUMP crystal showing diffractions at 2.2 Å; (b) the wtPL$_3$·dUMP crystal mounted for data collection, the size of the red square is 100µm×100µm.

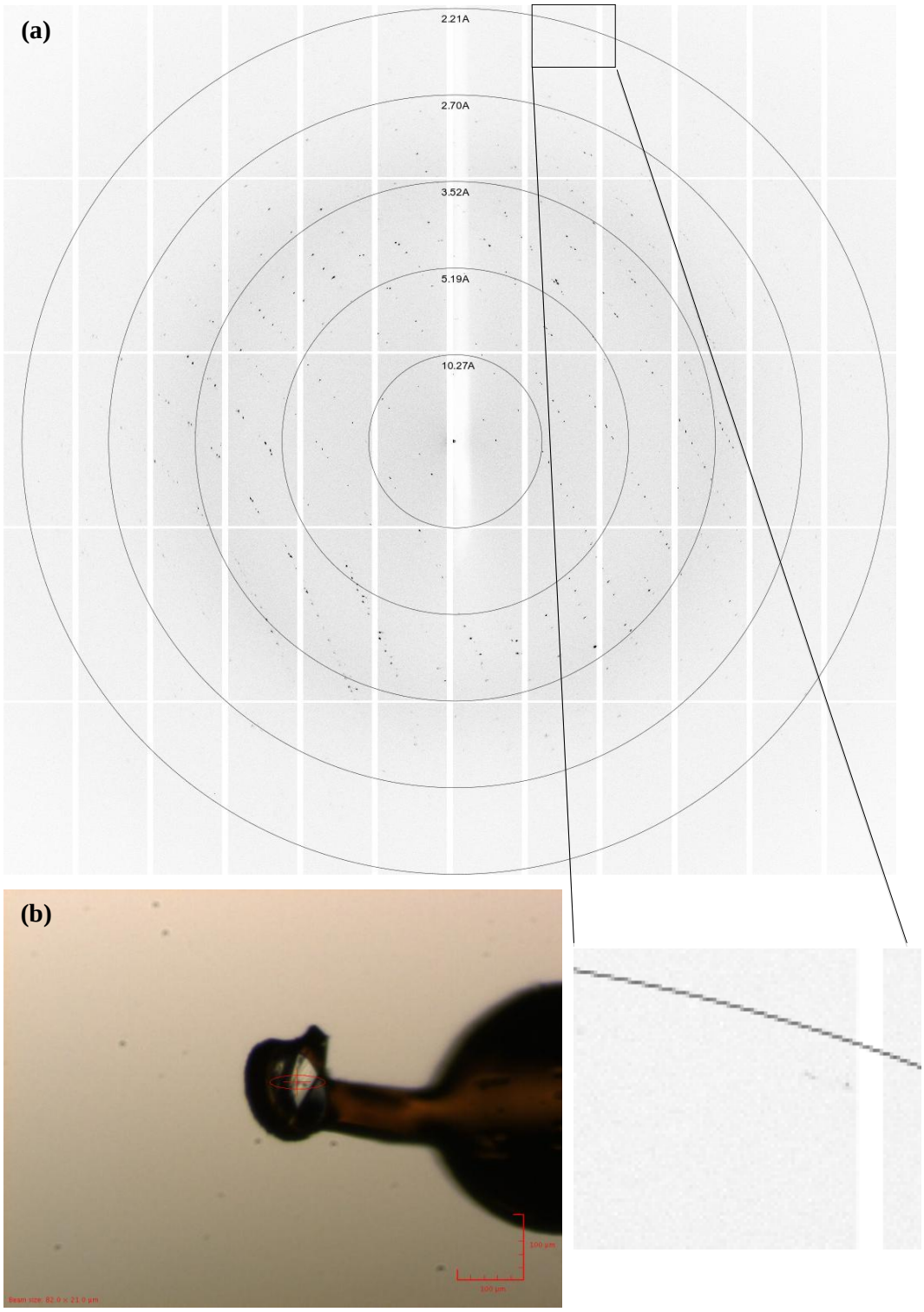


Figure 3.31. Crystals and X-ray diffraction picture of PL_3 D80A dCMP complex. (a) Diffraction image of a PL_3 D80A dCMP crystal showing diffractions at 2.25 Å; (b) The PL_3 D80A dCMP crystal mounted for data collection, the size of the red square is 100µm×100µm.

Table 3.17. Data collection and refinement statistics of wtPL$_3$ dUMP and PL$_3$ D80A dCMP.

	wtPL$_3$ dUMP	**PL$_3$ D80A dCMP**
A. Data Collection Statistics		
Space Group	$P2_12_12_1$	P1
Cell	a = 76.19 Å, b = 91.32 Å, c = 93.43 Å α = β = γ = 90 °	a = 30.96 Å, b = 62.51 Å, c = 70.22 Å α = 89.78 °, β = 79.21 °, γ = 78.90 °
No. of molecules in the ASU	4	4
Resolution Range (Å)	50 – 2.28	50 – 2.60
R_{sym} [a]	0.061 (0.467)*	0.103 (0.410)*
I/σ(I)	14.7 (2.2)*	27.0 (10.7)*
Completeness (%)	98.2 (99.2)*	91.5 (65.6)*
Redundancy	4.3 (4.3)*	3.1 (3.0)*
B. Refinement Statistics		
Resolution range (Å)	50 – 2.28	50 – 2.60
No. of reflections	126853	43228
Unique No. of reflections	29780	14168
R_{cryst} [b]	0.23	0.25
R_{free} [c]	0.29	0.30
Number of non-H atoms:		
Protein	3840	3864
Ligand	4 dUMP, 2 PO_4^{3-}	4 dCMP
Water	83	0
Average temperature factor (Å^2)	48.3	46.1
RMSD for bond lengths (Å)	0.01	0.01
RMSD for bond angles ( )	1.41	1.51

* Values in parentheses are for the outer shell, which is 2.34 – 2.28Å, and 2.72 – 2.60Å respectively.

[a] $R_{sym} = \frac{\sum_h \sum_i |I_i(h) - I(h)|}{\sum_h \sum_i I_i(h)}$, where $I_i(h)$ and $I(h)$ are the i^{th} and the mean measurements of the intensity of reflection *h*, respectively.

[b] $R_{cryst} = \frac{\sum_h |F_o - F_c|}{\sum_h F_o}$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection *h*, respectively.

[c] R_{free} is equal to R_{cryst} for a randomly selected 5% reflections not used in the refinement.

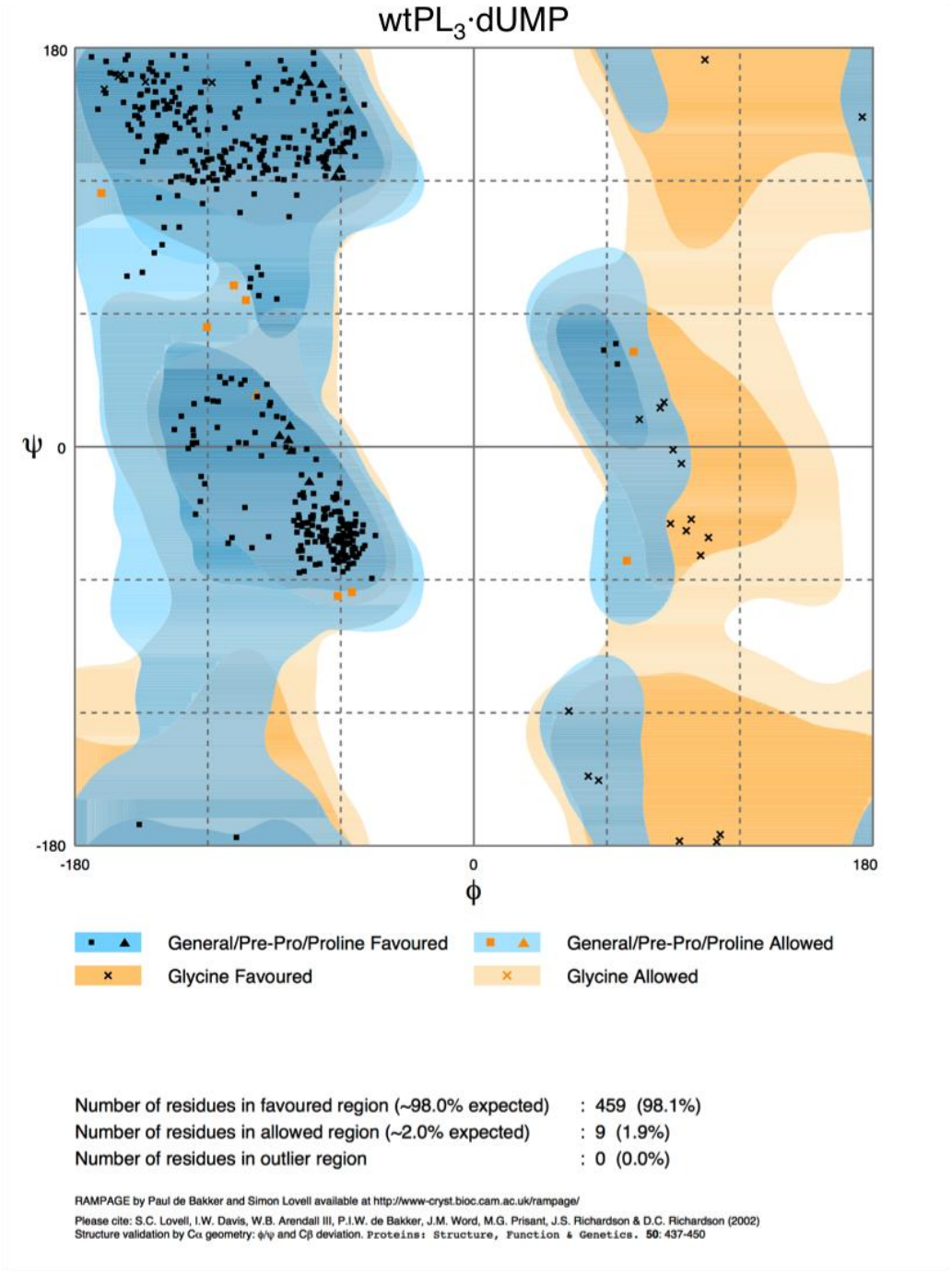


Figure 3.32. Ramachandran plot of wtPL$_3$·dUMP complex structure, which has four molecules in the ASU. 98.1% of the residues are lying in the most favoured regions. The Ramachandran plot was calculated using *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999).

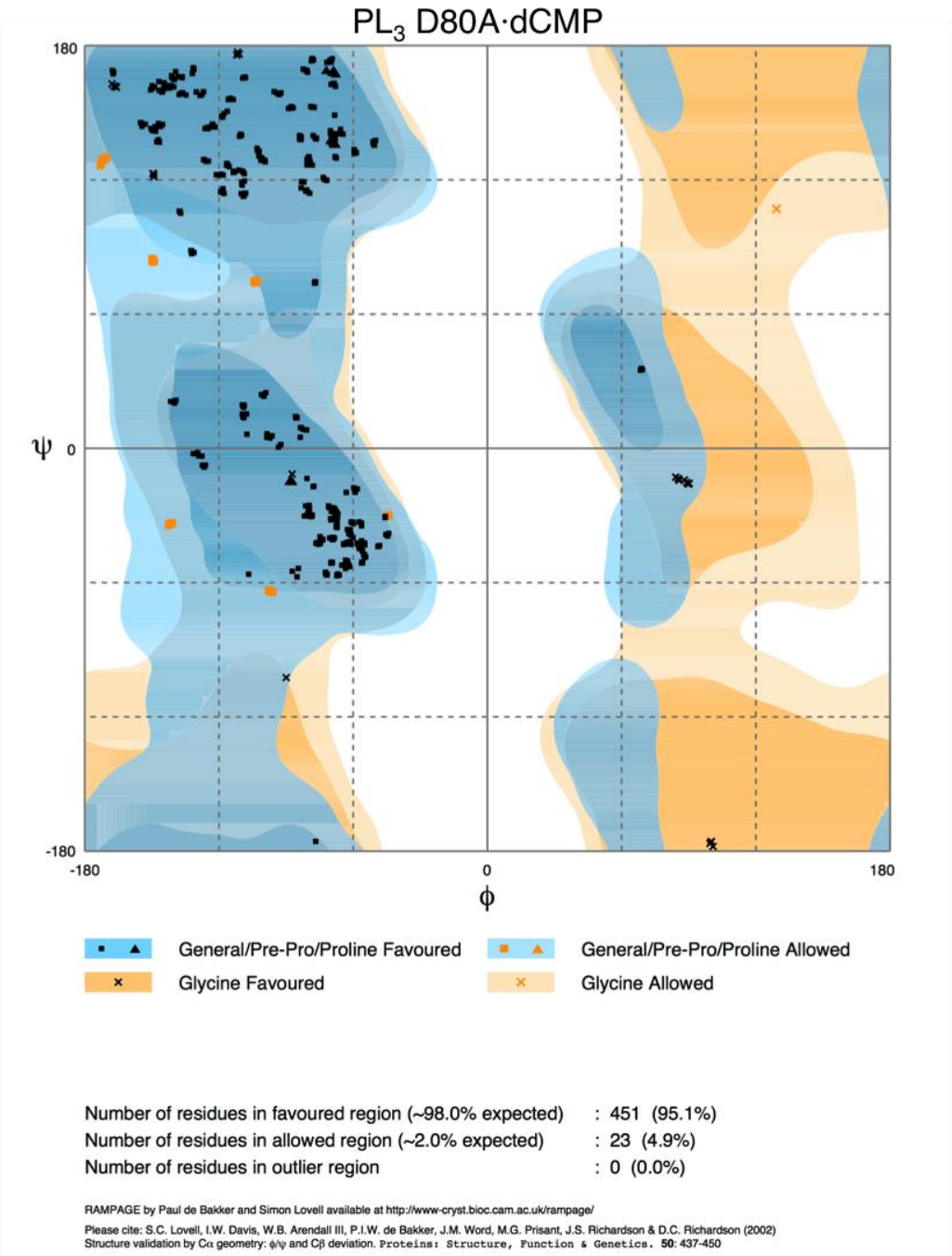


Figure 3.33. Ramachandran plot of the refined PL_3 D80A dCMP complex structure, which has four molecules in the ASU. 94.9% of the residues are lying in the most favoured regions. The Ramachandran plot was calculated using *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999).

3.4 Structural analysis

3.4.1 Structures of wtPL$_3$ and PL$_3$ D80A mutant

3.4.1.1 Overall structures

Structure of wtPL$_3$

The crystal structure of wtPL$_3$ was determined at 1.25 Å resolution with well defined electron density (Figure 3.34). There is one molecule in the ASU. The final model comprises residues -2 – 119 with 6 side chains in alternative conformations, 2 ethylene glycol molecules, and 130 water molecules. Ramachandran plot revealed that 96.7% of the residues in the structures are in the most favourable regions with the remainder in the allowed regions (Figure 3.28 in Section 3.3.15.1).

The structure of wtPL$_3$ adopts the characteristic α/β fold of RNase A, with three α helices and four β strands (Figure 3.35; Table 3.18). The structure is stabilised by four disulphide bonds (Figure 3.35). The main chain is well ordered apart from residues 84 – 86 in the L5 loop linking β2 and β3. High flexibility of corresponding regions is a characteristic feature of members of the RNase A superfamily, and is poorly defined in crystal structures of hEDN (Leonidas *et al.*, 2001; Mosimann *et al.*, 1996; Swaminathan *et al.*, 2002), hECP (Boix et al., 1999a), hANG (Leonidas *et al.*, 1999a), and murine ANG-1 (mANG-1) (Holloway *et al.*, 2005).

The RNA binding subsites are also identified in the structure of PL$_3$ (Figure 3.36). Subsite P$_0$ has only one residue, Lys-65; the catalytic subsite P$_1$ consists of Gln-11, His-12, Lys-40, and His-116; Arg-7 and Arg-10 contribute to subsite P$_2$. The subsite B$_1$ consists of residues Phe-43, Asn-43, Thr-44, and Phe-117, and has a preference for uracil base; residues Asn-66, Gln-68, Asn-70, and Glu-108 belong to the adenine-binding B$_2$ subsite.

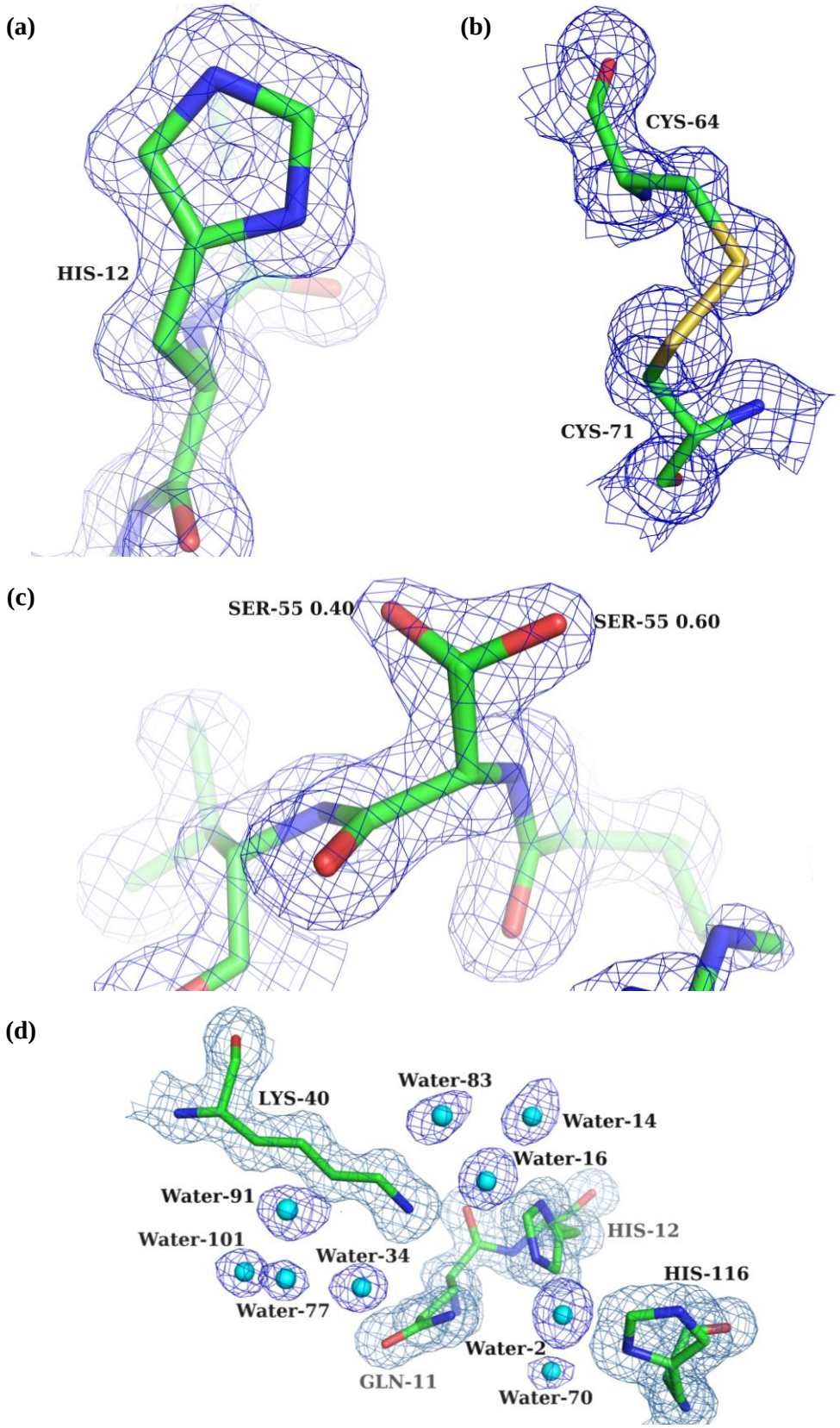


Figure 3.34. ***2F_o-F_c*** **electron density maps (contoured at 1.0σ) of the crystal structure of wtPL$_3$: (a) the catalytic histidine at position 12; (b) the disulphide bond formed by Cys-64 and Cys-71; (c) residue Ser-55 in alternative conformations; (d) the water network near the catalytic P$_1$ subsite.**

Table 3.18. Secondary structure elements of wtPL$_3$.

α helices	β strands	Loops
α1 Gln-1 – His-12	β1 Phe-42 – Ile-46	L1 Val-13 – Asn-21
α2 Asp-22 – Arg-31	β2 Cys-71 – Glu-83	L2 Arg-32 – Arg-41
α3 Ile-50 – Ile-56	β3 Tyr-94 – Glu-108	L3 His-47 – Asp-49
	β4 Val-113 – Phe-117	L4 Cys-57 – Asn-70
		L5 Thr-84 – Arg-93
		L6 Gly-109 – Glu-112
		L7 Asp-118 – Lys-119

All elements are defined by *PROCHECK* (Laskowski *et al.*, 1993; Vaguine *et al.*, 1999; Winn *et al.*, 2011); the three residues from the N-terminal tag are excluded.

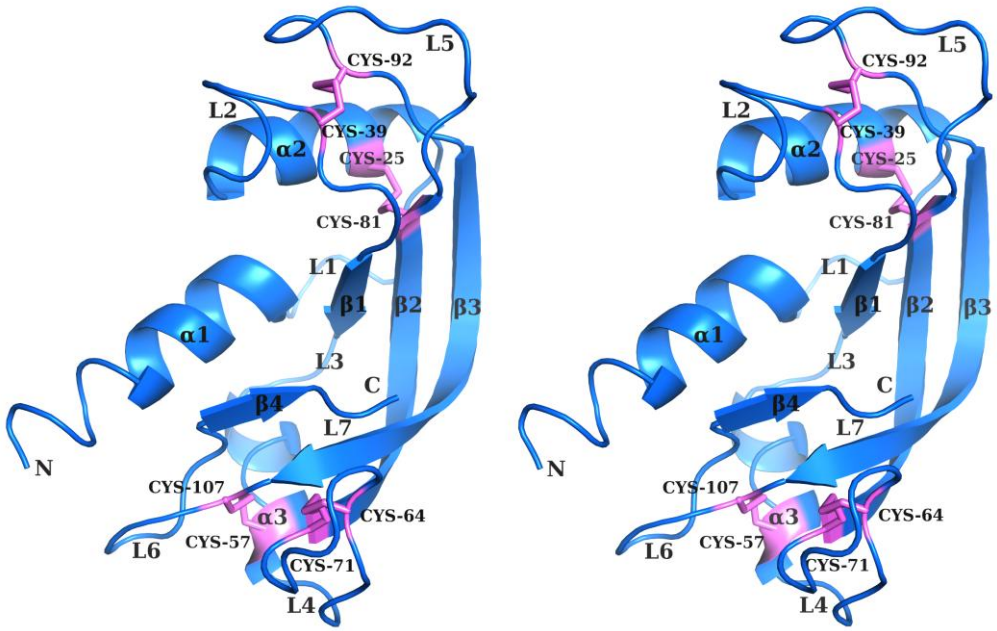


Figure 3.35. Cartoon representation of the structure of wtPL$_3$ in stereo view. Disulphide bond forming cystines are labelled and shown in violet; α helices, β strands and loop regions are labelled, the N- and C- termini are labelled as N and C respectively.

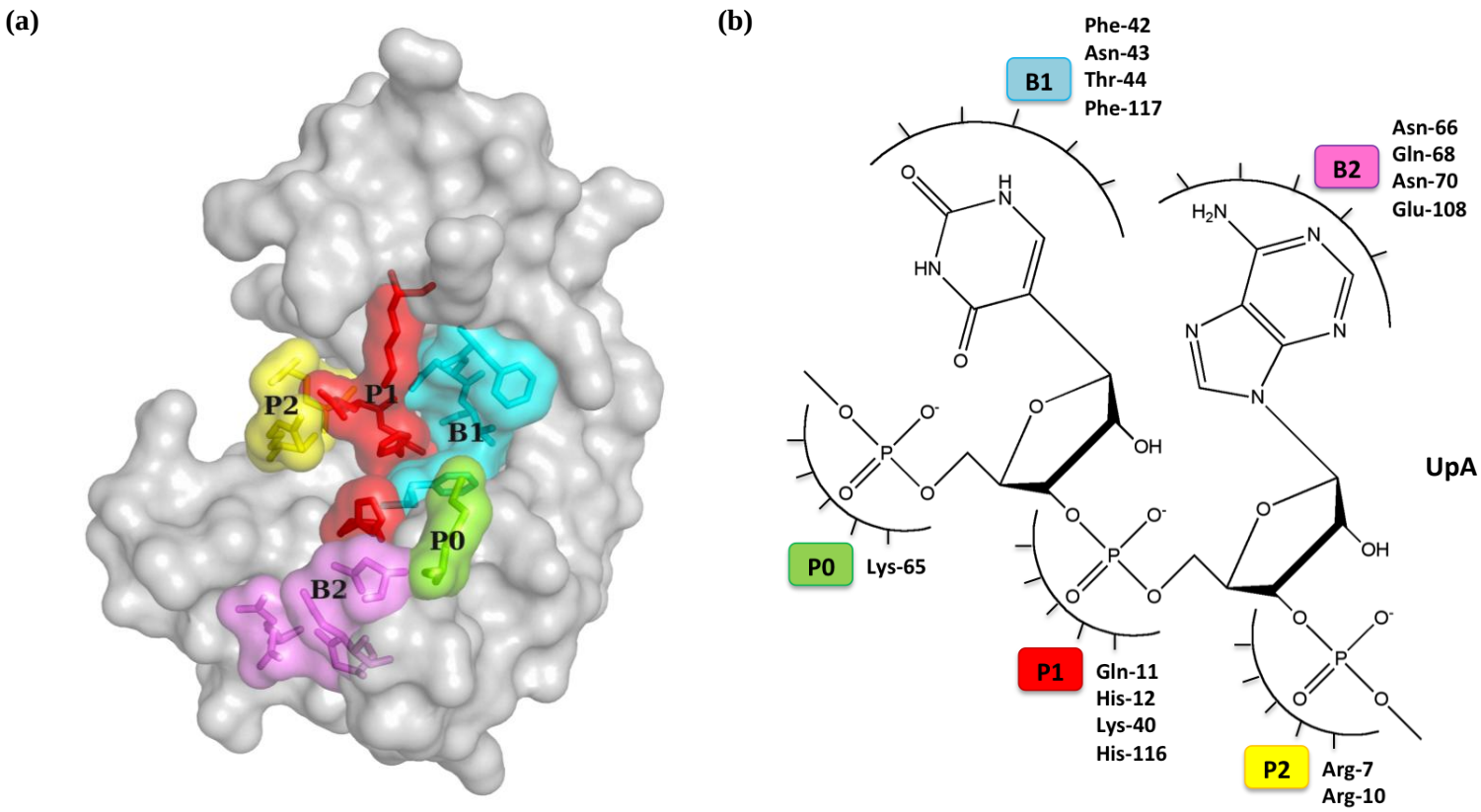


Figure 3.36. Subsites of PL_3 involved in the binding of RNA. (a) Surface diagram of the three-dimensional structure of PL_3 showing the subsites involved in the recognition and degradation of RNA. B and P indicate the base and phosphate binding subsites respectively; (b) Schematic representation of the recognition of an RNA fragment (UpA) by PL_3. Residues in the different subsites are labelled. The phosphate group of the phosphodiester bond to be hydrolysed by the enzyme binds to P_1; B_1 is specific for pyrimidines, B_2 shows a preference for purines.

Structure of PL_3 D80A mutant

The crystal structure of PL_3 D80A was determined at 2.9 Å resolution with four molecules in the ASU (Figure 3.37). The final model of PL_3 D80A includes residues 0 – 119 in molecule A, 1 – 119 in molecule B, -1 – 119 in molecules C and D. Ramachandran plot suggests that 95.1% of the residues in the structures of PL_3 D80A are in the most favourable regions with the remaining 4.9% in the allowed regions (Figure 3.29 in Section 3.3.15.1).

PISA analysis (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) (Krissinel and Henrick, 2007) of the structure suggests none of the potential dimers in the ASU appears to be biologically relevant. *PISA* server calculates a complexation significance score (CSS) based on the molecular contacts at the interface, indicating the likelihood for a packing interface observed in a crystal structure to be also observed in solution. Although there are one or two hydrogen bonds identified between A-D and B-C dimers (Table 3.19), the CSS scores for all the potential dimers are 0, suggesting the arrangement of the molecules in the ASU is a result of crystal packing.

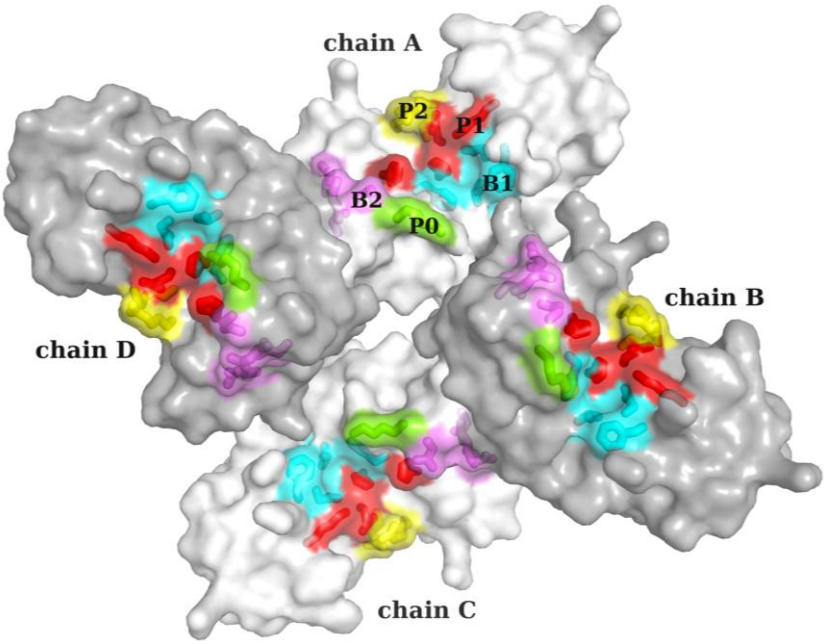


Figure 3.37. Arrangement of the four molecules of PL_3 D80A in the ASU. The P_0, P_1, P_2, B_1, and B_2 subsites are coloured in green, red, yellow, cyan, and pink.

Table 3.19. Potential hydrogen bonds at various dimer interfaces within the ASU of the PL_3 D80A structure.

Hydrogen bond (donor – acceptor)	Distance (Å)
DVal-75 N – A Gly-67 O	3.17
B Val-75 N – C Gly-67 O	3.05
C Gln-63 NE2 – B Glu-73 O	3.14

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

The four molecules of PL_3 D80A mutant in the ASU are highly similar: superposition of molecules B, C, and D with molecule A reveals an rmsd value of 0.1 Å, 0.04 Å, and 0.1 Å for C^{α} atoms respectively. The average B-factor of molecule C is the lowest amongst all, suggesting it is the most ordered molecule in the ASU. Hence it will be used as a reference for all further discussion. The position of Ala-80 is well defined in the electron density maps, an example is shown in Figure 3.38.

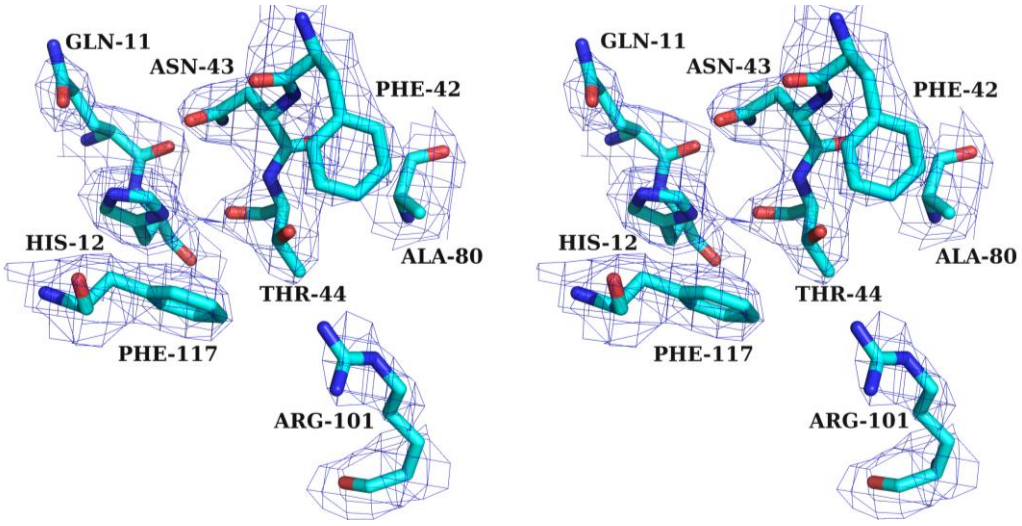


Figure 3.38. Stereo view showing the $2F_o$-F_c electron density map (contoured at 1.0σ) of the refined model of PL_3 D80A mutant in the region where the D80A mutation lies.

3.4.1.2 Comparison of PL_3 with hRNase 4

The comparison of $wtPL_3$ and hRNase 4 structures will be the main focus on here, as $wtPL_3$ has the most similar conformation as that of hRNase 4 within the RNase A superfamily. The levels of similarity between the structures of $wtPL_3$

and RNase A, hEDN, hECP, hRNase 4, and hANG were analysed by *DaliLite* (Holm and Park, 2000; McWilliam *et al.*, 2013) pairwise three-dimensional structure alignments.

Least-square superposition of $wtPL_3$ with hRNase 4 (Figure 3.39) yields an rms deviation of 0.398 Å for C^{α} atoms, indicating $wtPL_3$ is structurally very similar to hRNase 4. The secondary structure elements of $wtPL_3$ and hRNase 4 mostly coincide, with two minor differences in the lengths of α2 (Ile-50 – Ile-56 in $wtPL_3$ and Ile-50 – Ser-58 in hRNase 4) and β4 (Val-113 – Phe-117 in $wtPL_3$ and Val-113 – Gly-119 in hRNase 4). This is not at all surprising, as RNase 4 proteins are highly conserved, and $wtPL_3$ and hRNase 4 have a primary structure identity of 87.4%.

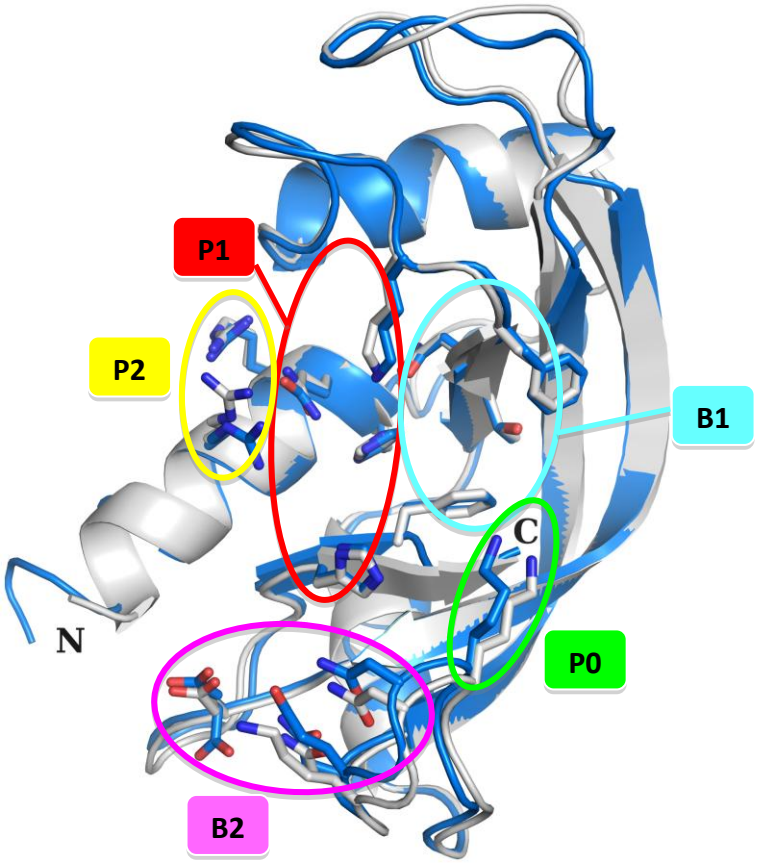


Figure 3.39. Superposition of the structure of PL_3 with that of hRNase 4 (PDB: 1RNF; Terzyan *et al.*, 1999). PL_3 and hRNase 4 are coloured in marine and grey respectively. Subsites P_0, P_1, P_2, B_1, and B_2 are marked by green, red, yellow, cyan, and pink circles; residues in these subsites are shown in stick representation.

The locations and residue compositions of the different RNA binding subsites of PL_3 are almost identical to those of hRNase 4 (Figure 3.39). Like hRNase 4, the conserved residues Gln-11, His-12, Lys-40, and His-116 comprise the P_1 catalytic site of PL_3; and side chain orientations of these residues are highly similar to those of the hRNase 4 (Figure 3.40).

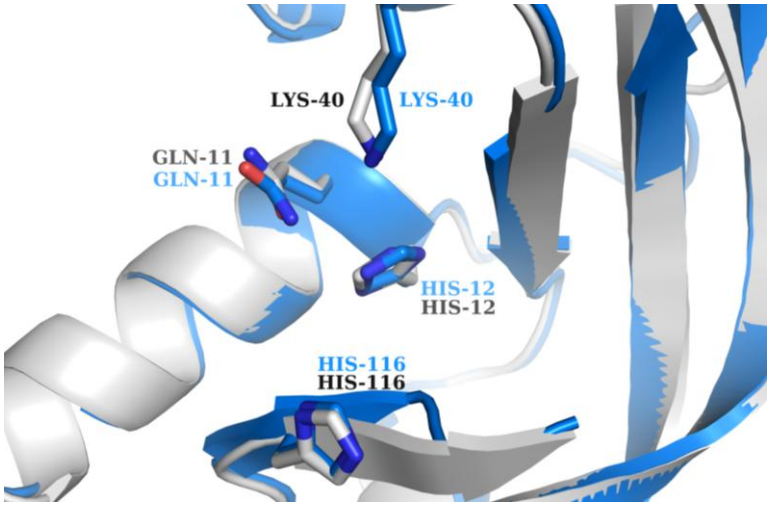


Figure 3.40. The P_1 active sites of PL_3 and hRNase 4 (PDB: 1RNF; Terzyan *et al.*, 1999). PL_3 and hRNase 4 are coloured in marine and grey respectively.

The P_0, P_2, B_1, and B_2 subsites are highly conserved between PL_3 and hRNase 4 (Figure 3.41-3.42); and the side chain orientations are highly similar. The less well superposed P_0 and B_2 subsites of PL_3 and hRNase 4 are likely due to the fact that they are located at loop regions with higher flexibility.

Amongst the four subsites, there is a single amino acid difference at position 68 in the B_2 subsite: residue lysine in hRNase 4 is substituted with glutamine in PL_3 (Figure 3.42b). B_2 subsite is where the recognition of the adenine base 3' to the phosphodiester bond being hydrolysed occurs, with one of the four residues being substituted, the tightness of the binding between the ligand and the protein could be slightly affected, which may in turn affects the turnover rate of RNA molecules. However, we still need coupled catalytic activities data of PL_3 and hRNase 4 to confirm such effect.

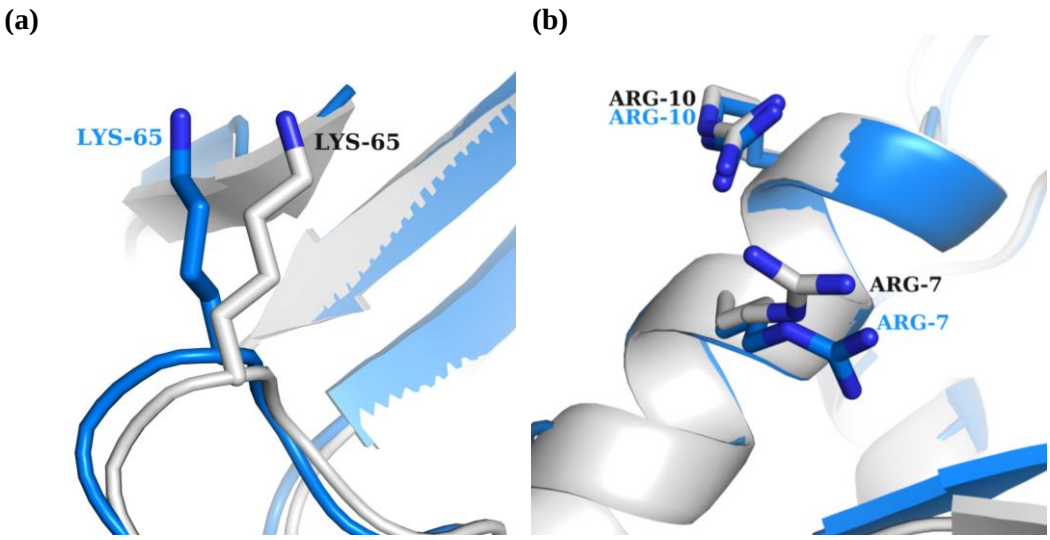


Figure 3.41. The (a) P_0 and (b) P_2 subsites of PL_3 and hRNase 4 (PDB: 1RNF; Terzyan *et al.*, 1999). PL_3 and hRNase 4 are coloured in marine and grey respectively.

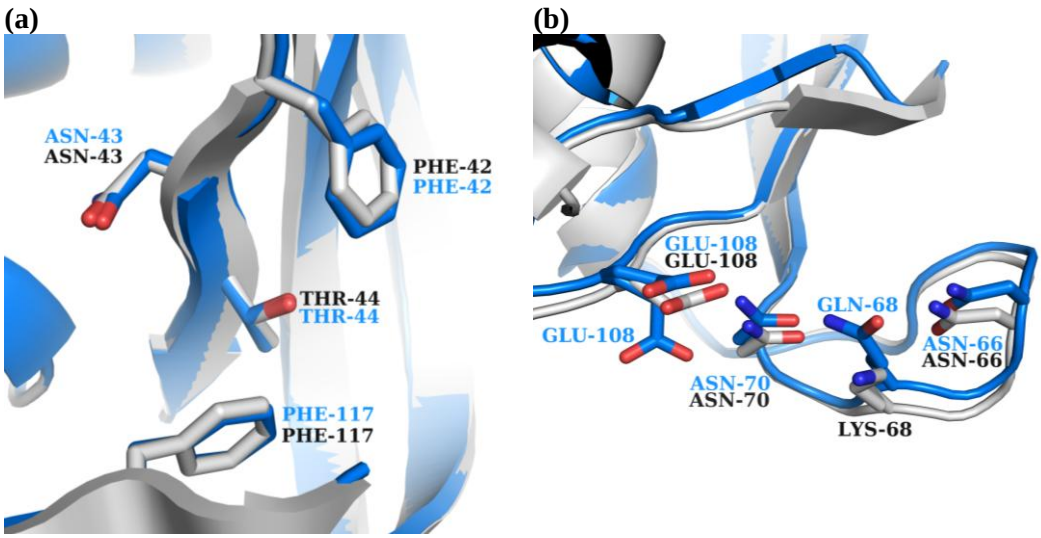


Figure 3.42. The (a) B_1 and (b) B_2 subsites of PL_3 and hRNase 4 (PDB: 1RNF; Terzyan *et al.*, 1999). PL_3 and hRNase 4 are coloured in marine and grey respectively. Glu-108 of PL_3 at the B_2 subsite is in two alternative conformations.

3.4.1.3 Comparison of PL_3 and the prototypic RNase A

Despite the different residue numbering, the catalytic residues Gln-11, His-12, Lys-40, and His-116 in the P_1 subsite of PL_3 are conserved in the structure of RNase A (Figure 3.43a). Such conservation is important for the ribonucleolytic activity of PL_3. The residue compositions of other subsites are highly similar in the structure of PL_3 and RNase A except for the B_1 pyrimidine-binding pocket (Figure 3.43b): Val-43 in RNase A is substituted with a phenylalanine residue

in PL_3 at position 42. Both site-directed mutagenesis studies on PL_3 (Vicentini *et al.*, 1996), and structural studies of hRNase 4 (Terzyan *et al.*, 1999) have revealed that Phe-42 along can change the cytidine/uridine specificity of RNase 4.

Another significant difference observed in the B_1 subsites of PL_3 and RNase A is that PL_3 lacks the corresponding residue of Ser-123 in RNase A. This early termination of peptide chain is also characterised in the structure of hRNase 4, and is considered to be involved in the substrate specificity of RNase 4 (Terzyan *et al.*, 1999).

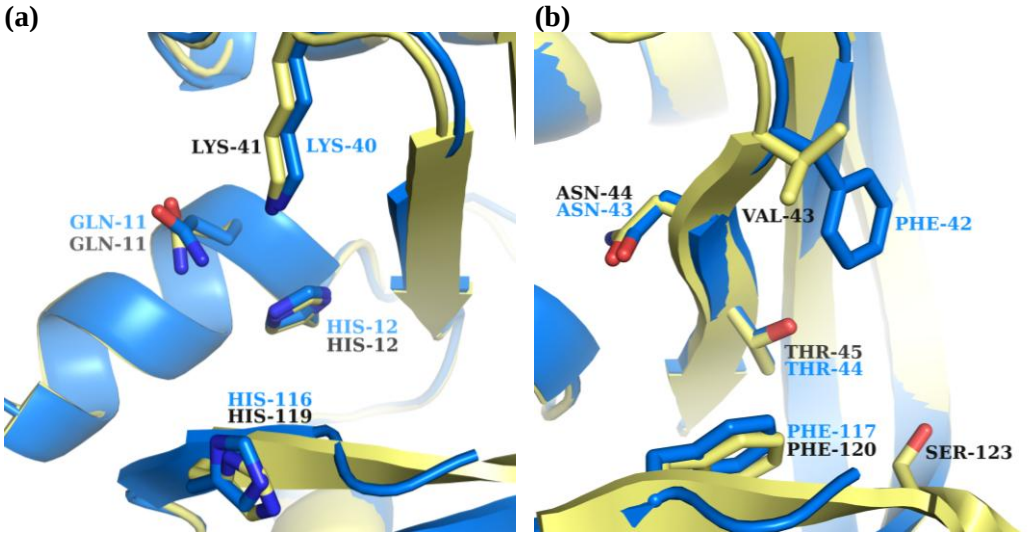


Figure 3.43. The (a) P_1 and (b) B_1 subsite of PL_3 and RNase A (PDB: 1AFU; Leonidas *et al.*, 1997). PL_3 and RNase A are coloured in marine and pale yellow respectively.

3.4.1.4 Difference in the interactions of wtPL_3 and PL_3 D80A mutant at the region of mutagenesis

On average, the Cα atoms of the mutant molecules deviate only 0.55 Å (rms) from those of PL_3, suggesting the aspartate to alanine substitution does not significantly affect the overall structure of the protein. Because the B_1 subsite of RNases determines their substrate preference, and the PL_3 D80A mutant is known to have different substrate specificity from PL_3, the interactions between residues located in the B_1 subsite and in the region of mutagenesis are analysed in detail.

In the structure of PL_3, 10 potential hydrogen bonds are identified involving either Asp-80 or residues in the B_1 subsite. 4 of the hydrogen bonds are side chain-side chain interactions, whereas the remainder involve main chain atoms (Figure 3.44, Table 3.20). The interactions involving the side chain of Asp-80 are lost in the structure of PL_3 D80A; the remaining hydrogen bonds identified in the structure of PL_3 are also observed in PL_3 D80A structure (Figure 3.44, Table 3.20).

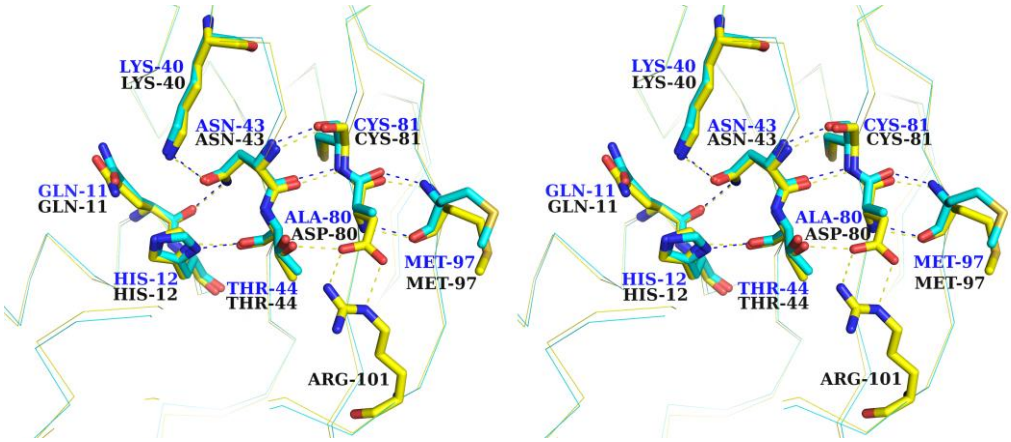


Figure 3.44. Comparison of the interactions at the B_1 subsite of PL_3 and PL_3 D80A mutant. Nitrogen, oxygen, and sulphur atoms are coloured in blue, red, and dark yellow respectively; carbon atoms of PL_3 and PL_3 D80A are coloured in yellow and cyan respectively. Residues of PL_3 are labelled in black, and residues of PL_3 D80A are labelled in blue. Potential hydrogen bonds are shown as yellow (PL_3) and blue (PL_3 D80A) dashed lines.

Table 3.20. Potential hydrogen bonds at the B_1 subsite of PL_3 and its D80A mutant.

Hydrogen bond (donor – acceptor)	Distance (Å)				
	PL_3	D80A			
		A	B	C	D
Asn-43 ND2 – Gln-11 O	2.88	2.76	2.77	2.75	2.75
His-12 ND1 – Thr-44 O	2.96	2.87	2.89	2.88	2.90
Lys-40 NZ – Asn-43 OD1	2.88	3.00	3.00	2.99	2.99
Asn-43 N – Cys-81 O	2.95	2.77	2.80	2.78	2.79
Cys-81 N – Asn-43 O	2.80	2.64	2.68	2.66	2.67
Thr-44 OG1 – Asp-80 OD1	2.69				
Arg-101 NH2 – Asp-80 OD1	2.92				
Arg-101 NE – Asp-80 OD2	2.92				
Asp-80 N – Met-97 O	2.90				
Met-97 N – Asp-80 O	2.87				
Ala-80 N – Met-97 O		2.77	2.78	2.78	2.79
Met-97 N – Ala-80 O		2.86	2.87	2.85	2.85

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown. Water mediated hydrogen bonds are not shown.

3.4.2 Structure of $wtPL_3$ dUMP complex

3.4.2.1 Overall structure of $wtPL_3$ dUMP complex

The crystal structure of $wtPL_3$ dUMP complex was determined at 2.28 Å resolution, with four molecules in the ASU (Figure 3.45). The final model includes 4 $wtPL_3$ molecules with residues 1 – 119, 4 dUMP molecules, 2 phosphate ions, and 83 water molecules. As suggested by the Ramachandran plot, all residues are either in the most favourable regions (98.1%) or in the allowed regions (1.9%) (Figure 3.32 in Section 3.3.15.2).

The CSS scores calculated by *PISA* server (Krissinel and Henrick, 2007) for all the potential dimers are 0 except those for molecules A-C (0.21) and B-C (0.03). With the highest CSS score (0.21) being much closer to 0 and the inconsistency in the residues involved in all potential dimerisations, the arrangement of the $wtPL_3$ dUMP molecules in the ASU is not biological. The interactions at various dimer interfaces will be discussed later.

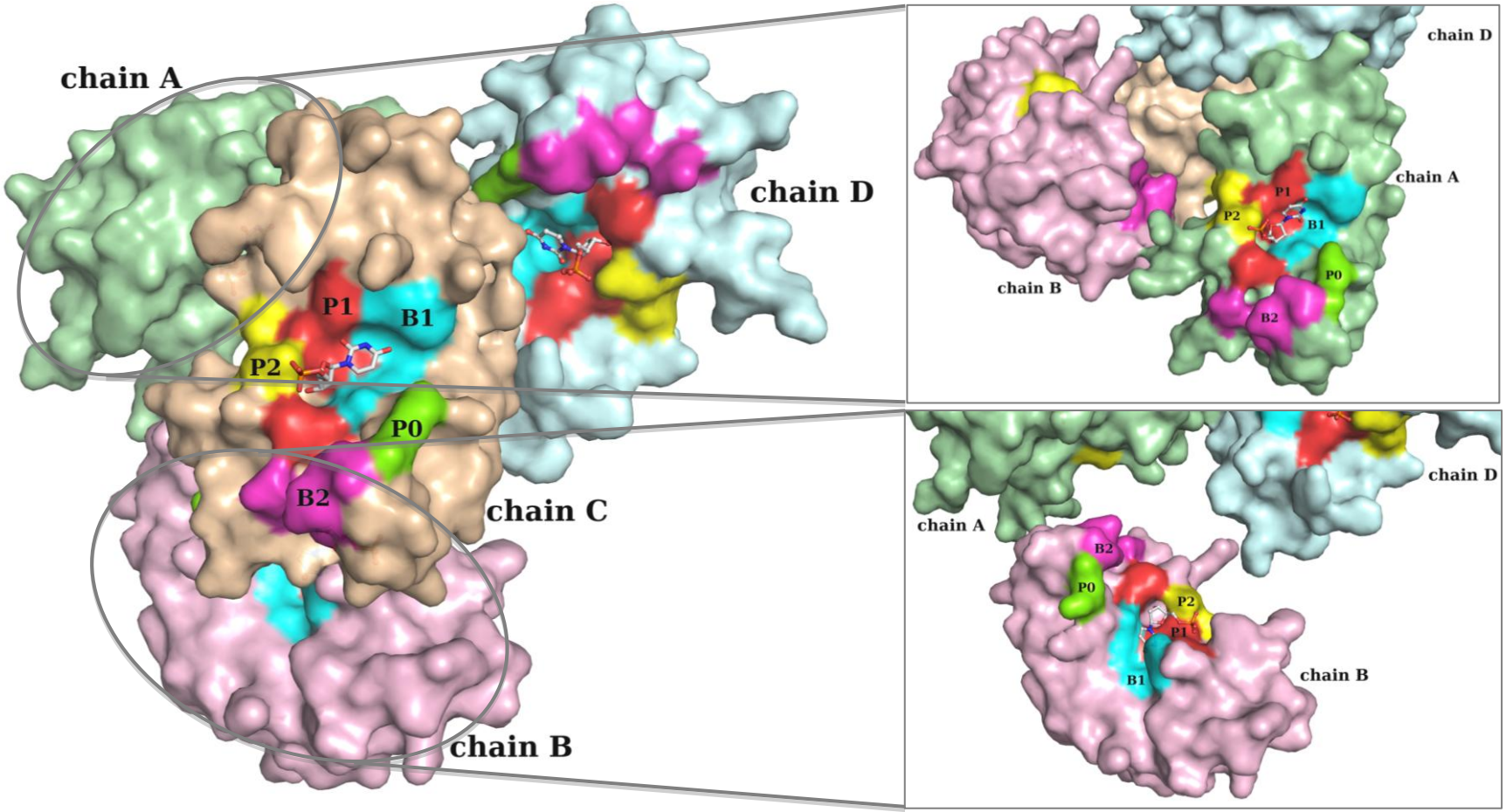


Figure 3.45. Arrangement of the four wtPL$_3$ dUMP molecules in the ASU. dUMP ligands are shown in stick representation; the P_0, P_1, P_2, B_1, and B_2 subsites are coloured in green, red, yellow, cyan, and pink.

3.4.2.2 dUMP recognition by PL_3

In the structure of PL_3 dUMP complex, the presences of the nucleotides at the P_1 catalytic sites of the PL_3 proteins were observed from the electron density maps (Figure 3.46). The phosphate groups of the four ligands have full occupancy (1.0) whereas the maximum occupancy for the nucleoside moieties is ~0.6, which are observed in dUMP-B and -C (bound to molecules B and C respectively). The difference in occupancy between different moieties of dUMP is possibly a result of the catalytic sites of the proteins being occupied partially by phosphate ions from the crystallisation solution, and partially by dUMP molecules.

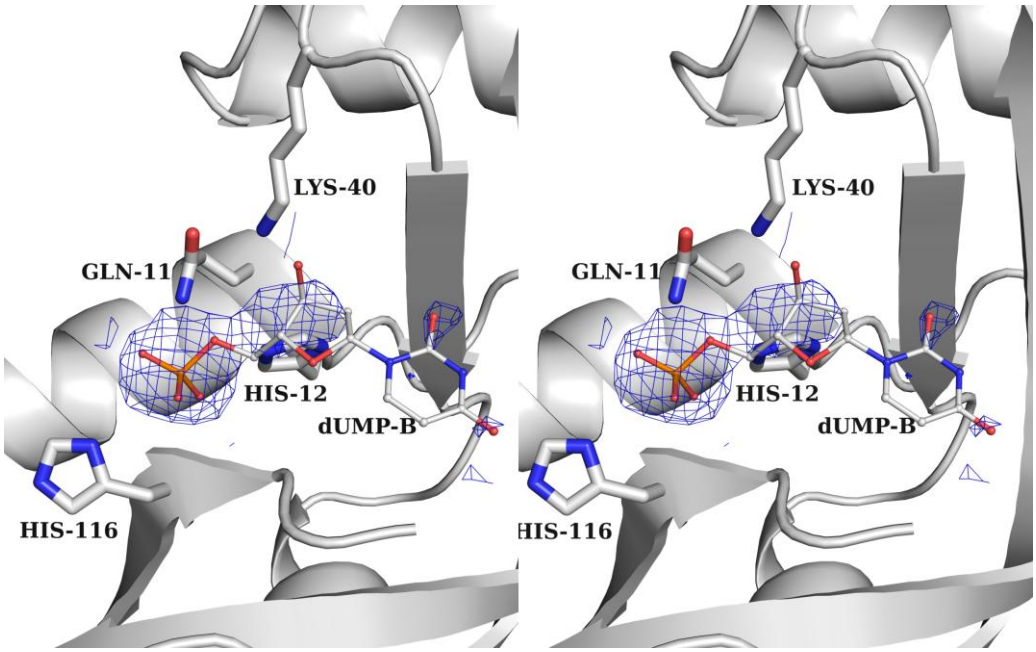


Figure 3.46. Close view of the P_1 catalytic subsite of PL_3 dUMP complex (chain B) in stereo view. The dUMP ligand is shown as stick model with $2F_o$-F_c electron density map contoured at 1.0σ.

In the structure of PL_3 dUMP complex, three of the ligands (dUMP-A to -C) adopt the C3'-*exo* conformation and only dUMP-D adopts the favoured C3'-*endo* conformation (Moodie and Thornton, 1993) (examples of sugar puckers are shown in Figure 3.47).

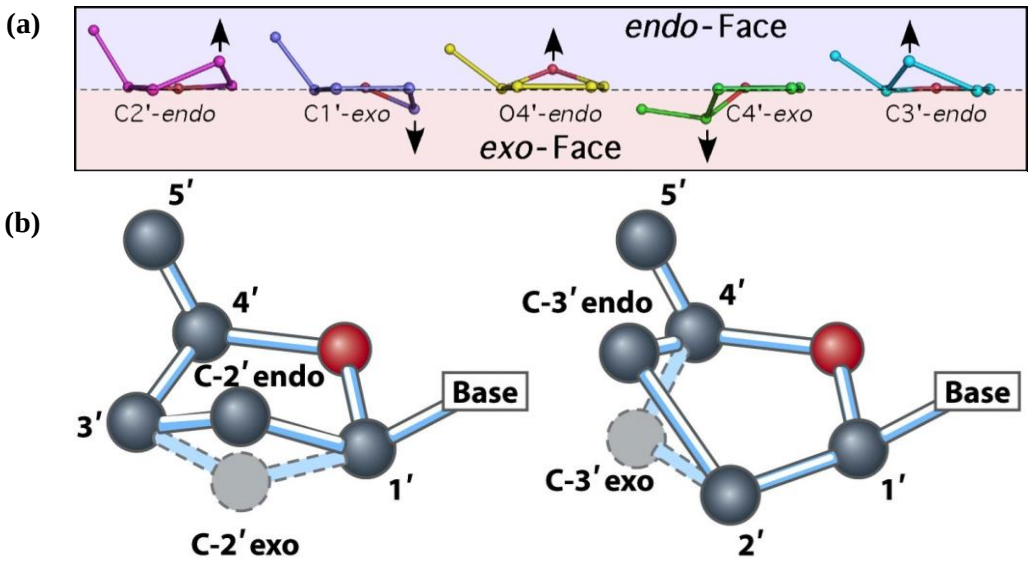


Figure 3.47. Conformations of nucleic acid sugar pucking. (a) Sugar ring pucker showing the *endo* (above) and *exo* (below) conformations of the 5-membered furanose sugar with the nucleotide base extended above the reference plane. Arrows indicate the atom that is puckered, and the direction of puckering. Figure was adapted from "DNA Structure: Alphabet Soup for the Cellular Soul" (Ho and Carter, 2011); (b) Examples of the C2'-*exo*/-*endo*, as well as C3'-*exo*/-*endo* conformations. Figure was adapted from "Lehninger Principles of Biochemistry, 5th Edition" (Nelson and Cox, 2008).

The four dUMP ligands interact with corresponding PL_3 molecules in a similar pattern: the uracil base is in the B_1 subsite, and the phosphate group locates in the P_1 subsite (Figure 3.48). Potential hydrogen bonds were observed between the uracil base and Thr-44 and Lys-119; the ribose and Gln-11, Lys-40, Asn-43, and Phe-117; and the phosphate with Arg-7, His-12, His-116, and Phe-117 (Figure 3.49-3.50, Table 3.21-3.22). None of the interactions observed in the four molecules are identical. The less well-defined electron density of the nucleoside moieties caused by their partial occupancies may contribute to such differences. Higher resolution data from crystals grown in the same condition or growing crystals in a PO_4^{3-} free crystallisation condition may improve the quality of the structure and provide us with more information on the interactions between the protein and its ligand.

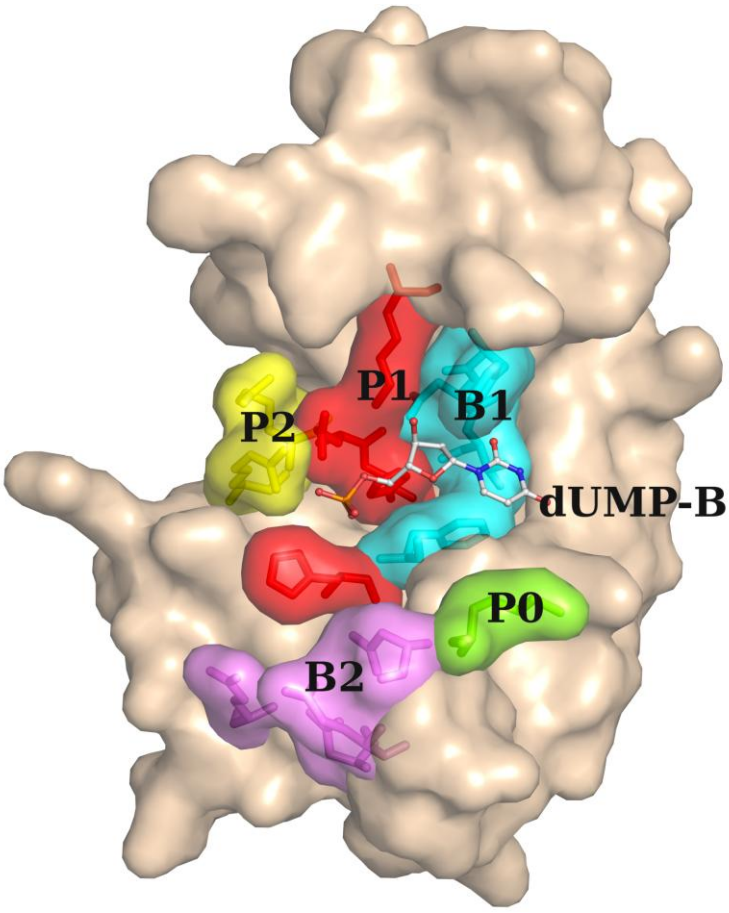


Figure 3.48. Surface diagram of the three-dimensional structure of PL_3 dUMP complex (chain B) with the subsites P_0 (Lys-65), P_1 (Gln-11, His-12, Lys-40, and His-116), P_2 (Arg-7 and Arg-10), B_1 (Phe-42, Asn-43, Thr-44, and Phe-117), and B_2 (Asn-66, Gln-68, Asn-70, and Glu-108) coloured in green, red, yellow, cyan, and pink respectively. Residues in the subsites as well as the dUMP-B ligand are shown in the stick representation. Carbon, nitrogen, oxygen, and phosphorus atoms of dUMP-B ligand are coloured in grey, blue, red and orange respectively.

The interaction between Arg-7 and the phosphate group is worth mentioning, as it is part of the P_2 subsite that should be interacting with the phosphate group of the second nucleotide (counting from 5' to 3', starting from the nucleotide binding to the B_1 and P_1 subsites) (Figure 3.36b). Similar interaction is also observed in the hRNase 4 dUp complex. A structure of PL_3 in complex with a dinucleotide, such as 2'-deoxyuridylyl-3', 5'-adenosine (2'-dUpA), may shed light on whether Arg-7 is actually part of the P_1 subsite in RNase 4 or such interactions only occur in the absence of the 3' nucleotide.

The crystal structure of PL_3 dUMP complex has a total of two PO_4^{3-} ions. One PO_4^{3-} interacts with the side chains of Arg-31 and Arg-32 of molecule B while the other one is incorporated at the interface of molecule A and C by a water molecule and the side chains of four arginine residues (Figure 3.51, Table 3.23). Since PO_4^{3-} ions are not required for the catalytic activity of PL_3 and no similar interactions were reported for the structures of other of the RNase A superfamily members, the presence of PO_4^{3-} ions in the crystal structure PL_3 dUMP complex is merely a crystallographic artefact.

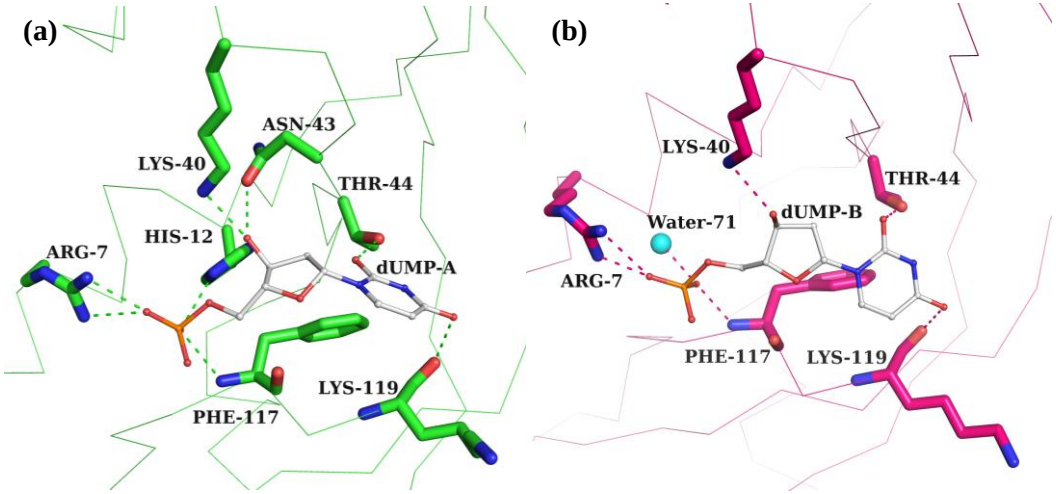


Figure 3.49. Details of the interactions at the catalytic site of PL_3 molecule and dUMP in (a) chain A and (b) chain B. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain A, chain B of PL_3 are coloured in green and pink; carbon atoms of dCMP-A and -B are coloured in grey. Water molecule is represented by sphere; potential hydrogen bonds are shown as dashed lines.

Table 3.21. Potential hydrogen bonds between dUMP-A and -B ligands and chain A and chain B of wtPL_3.

dUMP atom	PL_3 residue	Distance (Å)	dUMP atom	PL_3 residue	Distance (Å)
dUMP-A			dUMP-B		
O2	A Thr-44 OG1	2.65	O2	B Thr-44 OG1	3.02
O4	A Lys-119 O	2.95[a]	O4	B Lys-119 O	2.97[a]
O3'	A Lys-40 NZ	2.51	O3'	B Lys-40 NZ	2.96[a]
	A Asn-43 OD1	3.29[a]	OP1	B Arg-7 NH1	2.49
OP2	A Phe-117 N	3.00		B Arg-7 NH2	2.90
	A His-12 NE2	3.16[a]	OP2	B Phe-117 N	2.84
OP3	A Arg-7 NH1	3.02[a]		Water-71 O	2.71
	A Arg-7 NH2	2.65[a]			

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

[a] Potential interactions identified by *CONTACT* (Winn *et al.*, 2011).

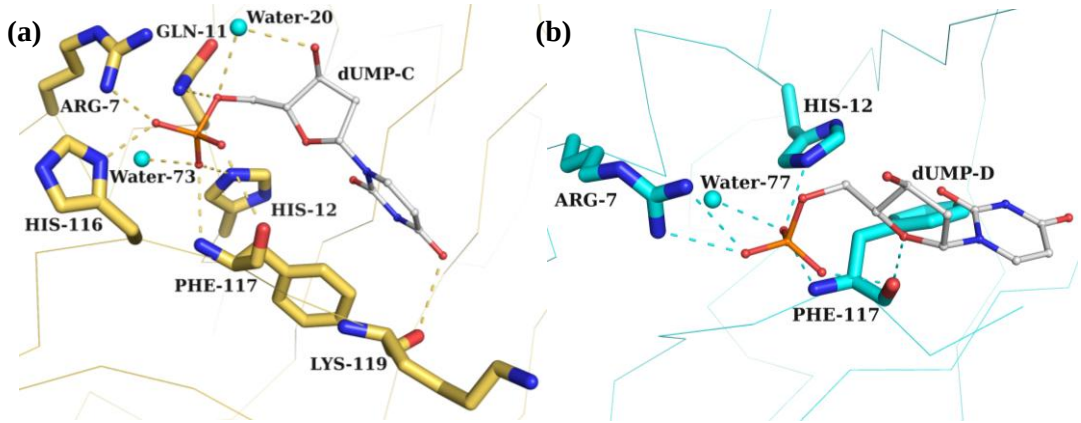


Figure 3.50. Details of the interactions at the catalytic site of PL_3 molecule and dUMP in (a) chain C and (b) chain D. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain C, chain D of PL_3 are coloured in yellow and cyan; carbon atoms of dCMP-C and -D are coloured in grey. Water molecule is represented by sphere; potential hydrogen bonds are shown as dashed lines.

Table 3.22. Potential hydrogen bonds between dUMP-C and -D ligands and chain C and chain D of wtPL_3.

dUMP atom	PL_3 residue	Distance (Å)	dUMP atom	PL_3 residue	Distance (Å)
dUMP-C			dUMP-D		
O4	C Lys-119 O	3.00[a]	O4'	D Phe-117 O	2.93[a]
O3'	Water-20 O	2.78	OP1	D Phe-117 O	3.22[a]
O5'	C Gln-11 NE2	3.28	OP2	D Arg-7 NH1	2.85
	Water-20 O	2.68		D Arg-7 NH2	3.18
OP1	C Arg-7 NH1	2.78	OP3	D His-12 NE2	2.85
	C His-116 ND1	2.71		D Phe-117 N	2.88
OP2	C His-12 NE2	2.77		Water-77 O	2.84
	C Phe-117 N	2.95			
	Water-73 O	2.70			
OP3	C Phe-117 O	3.24[a]			

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 ° identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

[a] Potential interactions identified by *CONTACT* (Winn *et al.*, 2011).

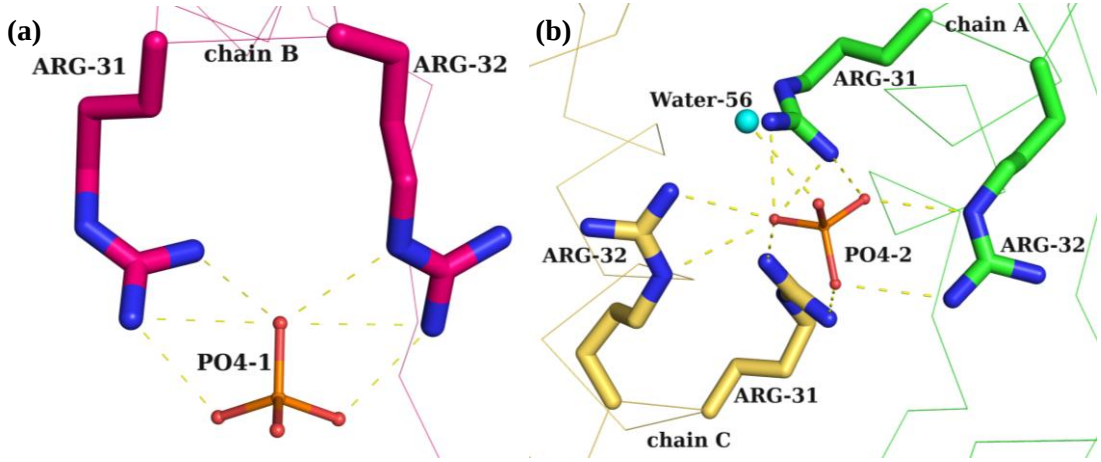


Figure 3.51. Details of the interactions between phosphate ion and the PL_3 molecule in (a) chain B and (b) chain A and C. Nitrogen, oxygen, and phosphorus atoms are coloured in blue, red and orange respectively; carbon atoms of the chain A, chain B, chain C of PL_3 are coloured in green, pink and yellow respectively. Water molecule is represented by sphere; potential hydrogen bonds are shown as dashed lines.

Table 3.23. Potential hydrogen bonds between PO_4^{3-}-1 and the chain B of the wtPL_3 protein, as well as PO_4^{3-}-2 and the chain A and C of the wtPL_3 protein.

PO_4^{3-}-1 atom	PL_3 residue	Distance (Å)	PO_4^{3-}-2 atom	PL_3 residue	Distance (Å)
O1	B Arg-32 NH2	2.76	O1	Water-56 O	2.73
O2	B Arg-31 NH1	2.74	O2	A Arg-32 NH2	3.16
	B Arg-31 NH2	3.04		C Arg-31 NH1	3.17
	B Arg-32 NE	2.91	O3	A Arg-31 NH2	2.74
	B Arg-32 NH2	3.29		A Arg-31 NH1	3.23[a]
O3	B Arg-31 NH2	2.80		C Arg-31 NH2	2.33
				C Arg-32 NE	3.27
				C Arg-32 NH2	2.92
			O4	A Arg-31 NH1	2.70
				A Arg-32 NE	2.84

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

3.4.2.3 *Comparison of ligand recognition by PL_3 with that by hRNase4 and RNase A*

Comparisons between PL_3 dUMP and hRNase 4 dUp complexes

Terzyan *et al.* (1999) reported a complex structure of hRNase 4 dUp (2'-deoxyuridine 3'-phosphate) along with the free structure of hRNase 4 to shed light on the structural basis for the substrate specificity. In the structure of hRNase 4 dUp complex, Arg-7, His-12, Thr-44, Phe-117 are involved in hydrogen bonding with dUp either directly or via a water network (Figure 3.52, 3.53). Potential hydrogen bonds are observed between the same residues in PL_3 and dUMP; however, different atoms on dUMP are involved in the interactions. The reason for the difference could lie in the different positions of the phosphate group of dUp and dUMP (Figure 3.52, 3.53).

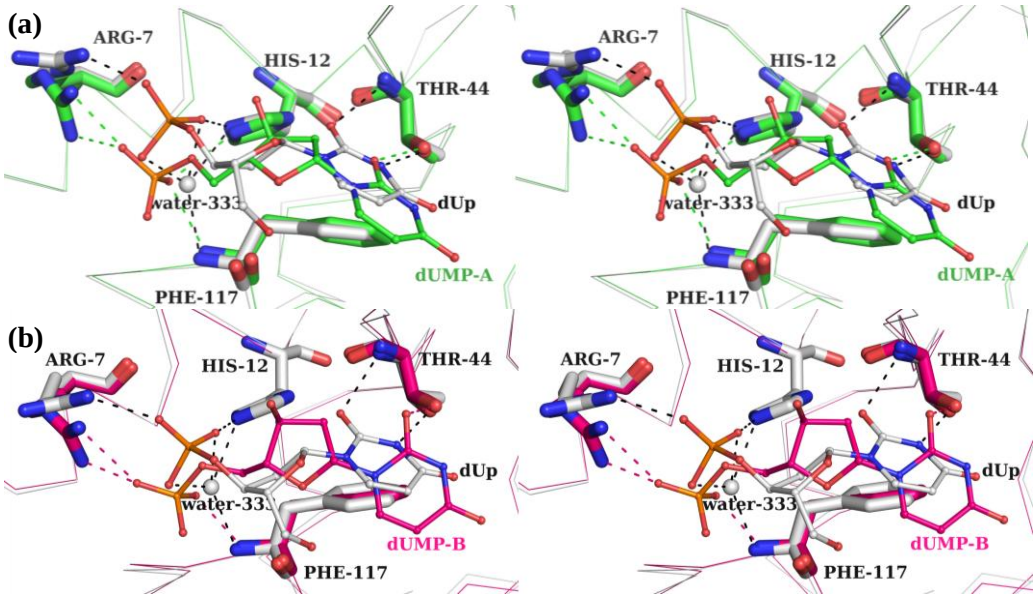


Figure 3.52. In the structures of hRNase 4 dUp (PDB: 2RNF; Terzyan *et al.*, 1999) and PL_3 dUMP complexes, residues Arg-7, His-12, Thr-44, and Phe-117 of PL_3 interact with the ligand in different ways. (a) Comparison between the interactions in the hRNase 4 dUp complex and chain A of the PL_3 dUMP complex; (b) comparison between the interactions in the hRNase 4 dUp complex and chain B of the PL_3 dUMP complex. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms in the structures of hRNase 4 dUp, chain A and chain B of PL_3 dUMP complexes are coloured in grey, green and pink respectively. Potential hydrogen bonds are shown as black dashed lines for hRNase 4 dUp complex, green and pink dashed lines for chain A and chain B of the PL_3 dUMP complex.

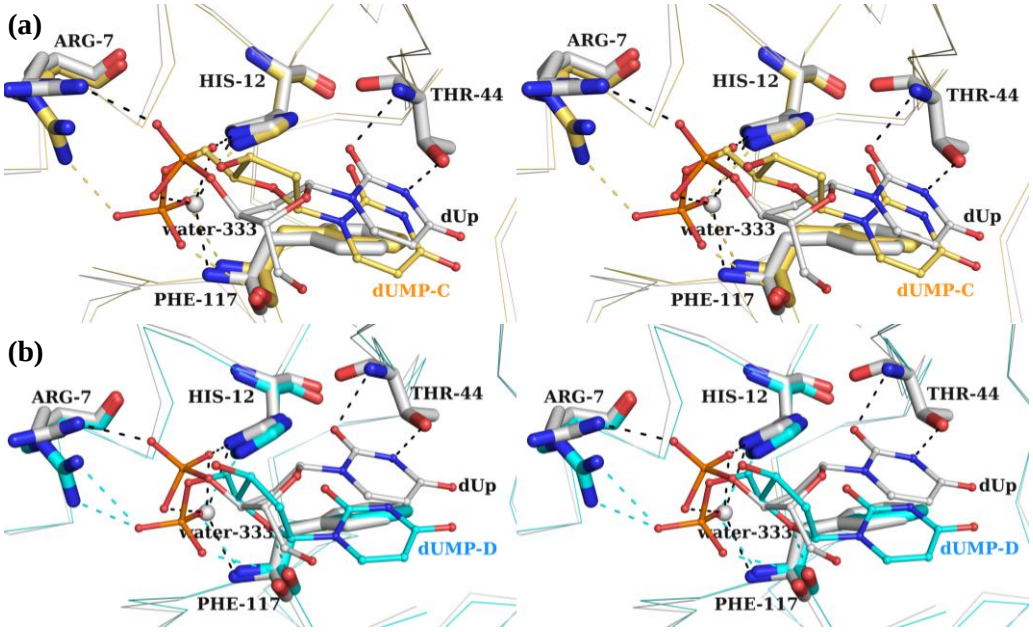


Figure 3.53. In the structures of hRNase 4 dUp (PDB: 2RNF; Terzyan *et al.*, 1999) and PL_3 dUMP complexes, residues Arg-7, His-12, Thr-44, and Phe-117 of PL_3 interact with the ligand in different ways. (a) Comparison between the interactions in the hRNase 4 dUp complex and chain C of the PL_3 dUMP complex; (b) comparison between the interactions in the hRNase 4 dUp complex and chain D of the PL_3 dUMP complex. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms in the structures of hRNase 4 dUp, chain C and chain D of PL_3 dUMP complexes are coloured in grey, yellow and cyan respectively. Potential hydrogen bonds are shown as black dashed lines for hRNase 4 dUp complex, yellow and cyan dashed lines for chain C and chain D of the PL_3 dUMP complex respectively.

The interaction between Arg-101 and the dUp observed in the hRNase 4 dUp structure, however, is not observed in the structure of PL_3 dUMP: although the O4 atom of dUMP is pointing towards the side chain of Arg-101, they are distanced too far apart (> 4 Å) for a hydrogen bond to form (Figure 3.54a, 3.54b). The O4 atom of dUMP is interacting with the main chain oxygen of Lys-119 of PL_3 to stabilise the ligand in the binding pocket (Figure 3.54c).

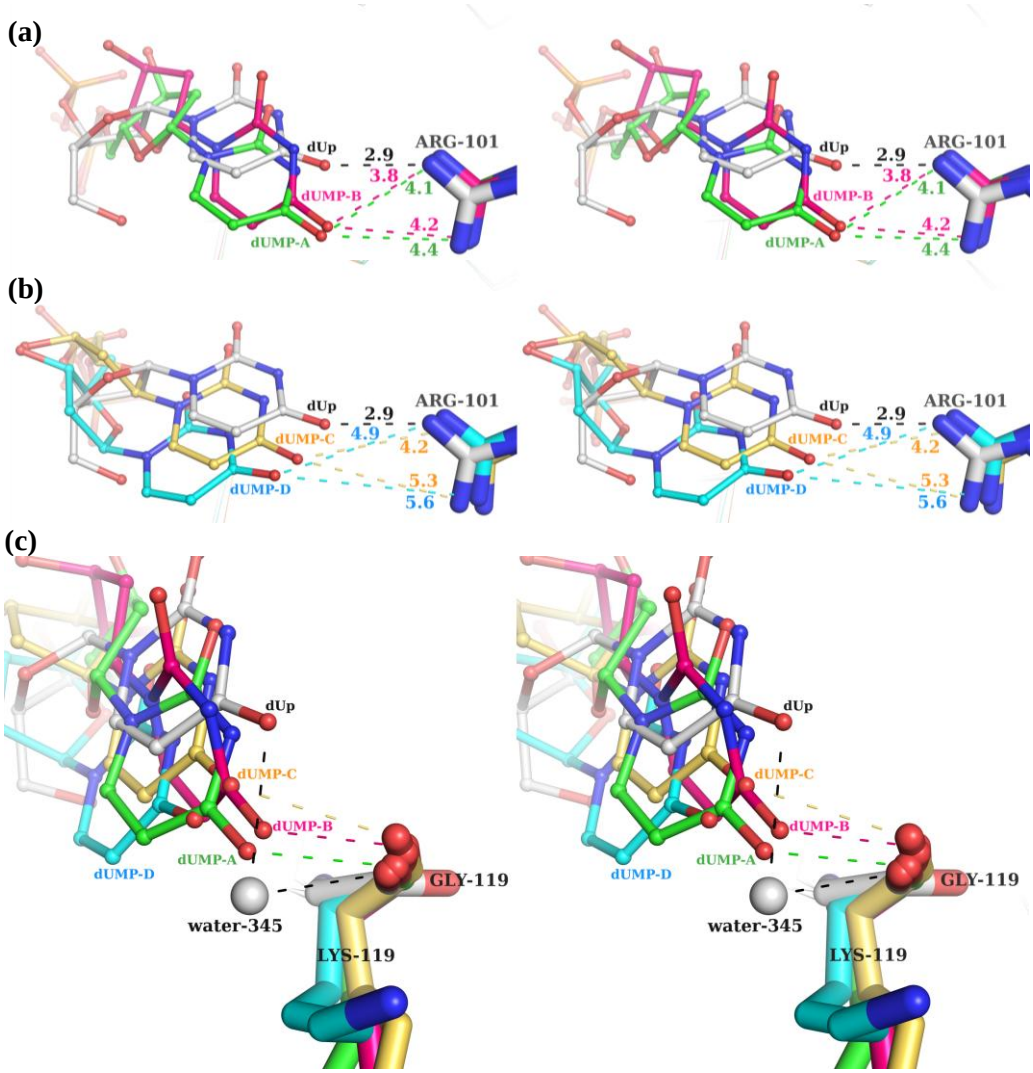


Figure 3.54. In the structures of hRNase 4 dUp (PDB: 2RNF; Terzyan *et al.*, 1999) and PL_3 dUMP complexes, the dUp/dUMP ligands are held by different residues. (a) Arg-101 of hRNase 4 forms a hydrogen bond with the dUp ligand, whereas Arg-101 in chain A and B of the PL_3 dUMP complex are distanced too far apart from dUMP for a hydrogen bond to form; (b) Arg-101 in chain C and D of the PL_3 dUMP complex are also too far from dUMP to form hydrogen bonds; (c) dUp interacts with Gly-119 of hRNase 4 via H_2O-345, where as dUMP ligands in chain A, B, and C of PL_3 interact with Lys-119 directly via hydrogen bonds. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms in the structures of hRNase 4 dUp, chain A, B, C and D of PL_3 dUMP complexes are coloured in grey, green, pink, yellow and cyan respectively. Potential hydrogen bonds are shown as black dashed lines for hRNase 4 dUp complex, green, pink, yellow and cyan dashed lines for chain A, B, C and D of the PL_3 dUMP complex respectively.

Terzyan *et al.* (1999) proposed that the bulky side chain of Phe-42 of hRNase 4 forces Asp-80 to hydrogen bond with Thr-44, which in turn interacts with the

N3 atom of the dUp molecule. The same side chain conformation of Phe-42 is observed in the structures of free PL_3, and molecules A, C, and D of PL_3 dUMP. Interestingly, the side chain of Phe-42 in molecule B points away from Asp-80, presenting a different conformation, and yet dUMP is still bound to the B_1 active site of molecule B (Figure 3.55). Since the B_1 subsite of molecule B is at the interface of molecules B and C, it is likely that such packing of molecules helps molecule B to hold dUMP-B in its B_1 subsite even when Phe-42 is in a different conformation.

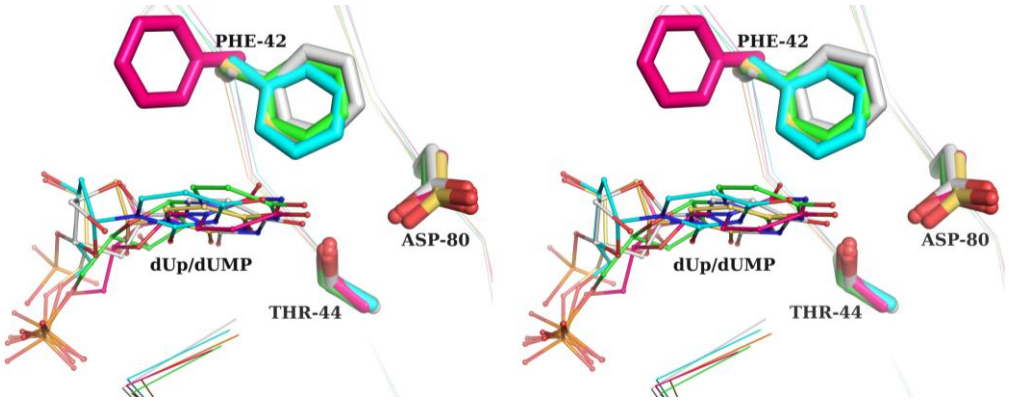


Figure 3.55. The different orientation of Phe-42 in chain B (pink) of the PL_3 dUMP complex compared with those in the hRNase 4 dUp complex (grey) (PDB: 2RNF; Terzyan *et al.*, 1999) as well as chain A (green), C (yellow), and D (cyan) of the PL_3 dUMP complex.

The effect of crystal packing on the conformation of Phe-42 of the four molecules was examined in more details. Phe-42 of molecules B and D are at the interface with different regions of molecule C within the ASU (Figure 3.56a). In molecule B, the side chain of Phe-42 is held away from Asp-80 (Figure 3.55) by the side chain carbon atoms of Arg-41 (of molecule B) via van der Waals interactions (Figure 3.56b). Unlike Arg-41 of molecules A, C, and D which are pointing away from Phe-42 (Figure 3.56c, 3.57b, 3.58b), the side chain of Arg-41 in molecule B forms hydrogen bonds with Cys-57 and Glu-112 in molecule C, thus folds towards and interacts with Phe-42 (Figure 3.56b). At the interface between molecules C and D, however, Phe-42 of molecule D is free from the attractions of Arg-41 of molecule D, which hydrogen bonded with Thr-18, Val-78, and Ser-99 of molecule C (Figure 3.56c).

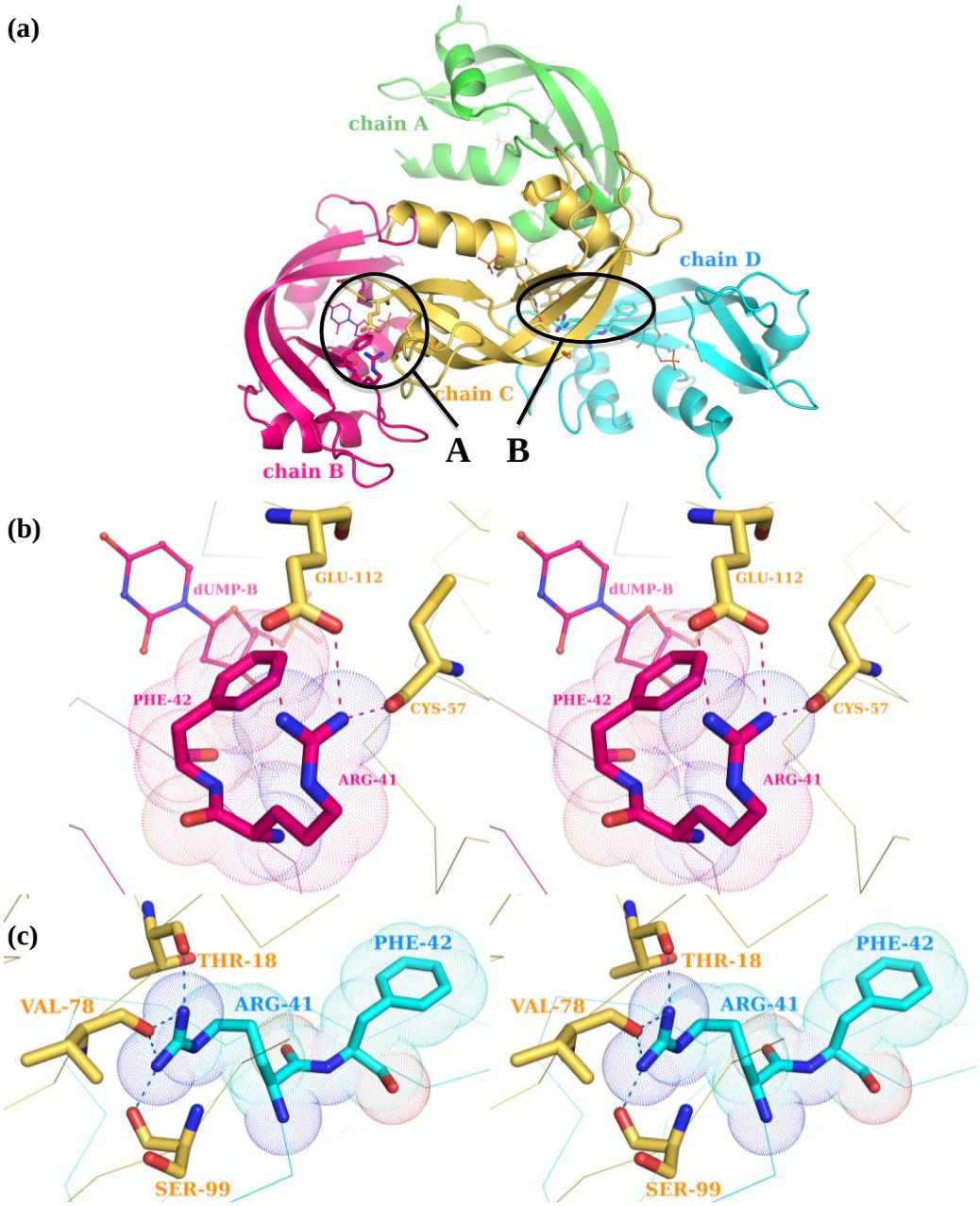


Figure 3.56. Within one ASU of the PL_3 dUMP complex, the conformation of Phe-42 in chain B is affected by crystal packing whereas the conformation of Phe-42 in chain D is not. (a) The four molecules of the PL_3 dUMP complex in the ASU with the interfaces involving Phe-42 in chain B and chain D marked as "A" and "B" respectively; (b) Close view of the interface near Phe-42 in chain B (pink) and chain C (yellow) of the PL_3 dUMP complex. Cys-57 and Glu-112 in chain C form hydrogen bonds with Arg-41 in chain B, which in turn attracts the side chain of Phe-42 and holds it in a different conformation; (c) Close view of the interface near Phe-42 in chain D (cyan) and chain C (yellow) of the PL_3 dUMP complex. In chain D, the side chain of Arg-41 faces away from Phe-42, and hydrogen bonds with Thr-18, Val-78, and Ser-99 in chain C. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms of chain B, C and chain D of the PL_3 dUMP complex are coloured in pink, yellow and cyan respectively.

Phe-42 of molecules A and C are involved in the interfaces with molecules of other ASUs (Figure 3.57a, 3.58a). At the interface between molecule A and molecule C of another ASU (Figure 3.58a), neither Phe-42 nor Arg-41 of molecule A is in enough close proximity to interact with residues (Cys-57 to Thr-59, Gly-109, and Asn-110) of molecule C (Figure 3.58b).

Similarly, at the interface between molecule C and molecule B of another ASU (Figure 3.58a), Arg-41 and Phe-42 of molecule C are too distanced from Glu-73 to Val-76 of molecule B (Figure 3.58b). In both cases, the conformation of Phe-42 is not affected by crystal packing.

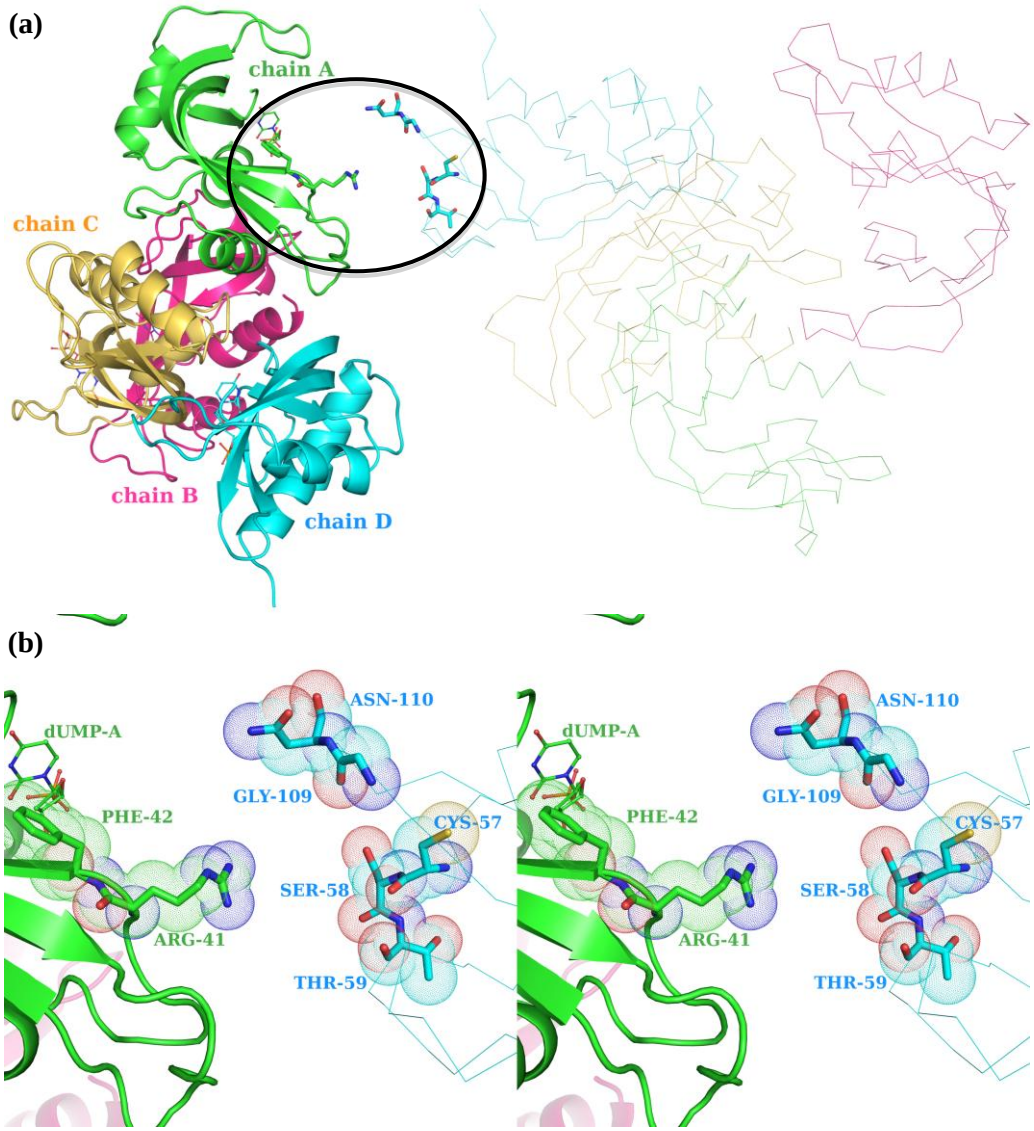


Figure 3.57. The conformation of Phe-42 in chain A of the PL_3 dUMP complex is not affected by crystal packing. (a) The four molecules of the PL_3 dUMP complex in the ASU showing in cartoon form with one of the adjacent ASU (showing in ribbon form), whose chain C is at the interface with chain A of the original ASU; (b) Close view of the interface near Arg-41 and Phe-42 in chain A of the PL_3 dUMP complex (green) and chain C in the adjacent ASU (cyan). Cys-57, Ser-58, Thr-59, Gly-109, and Asn-110 in chain C are near the interface; however, they are too distanced apart to interact with Arg-41 or Phe-42 in chain A. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms of chain A and chain D of the PL_3 dUMP complex are coloured in green and cyan respectively.

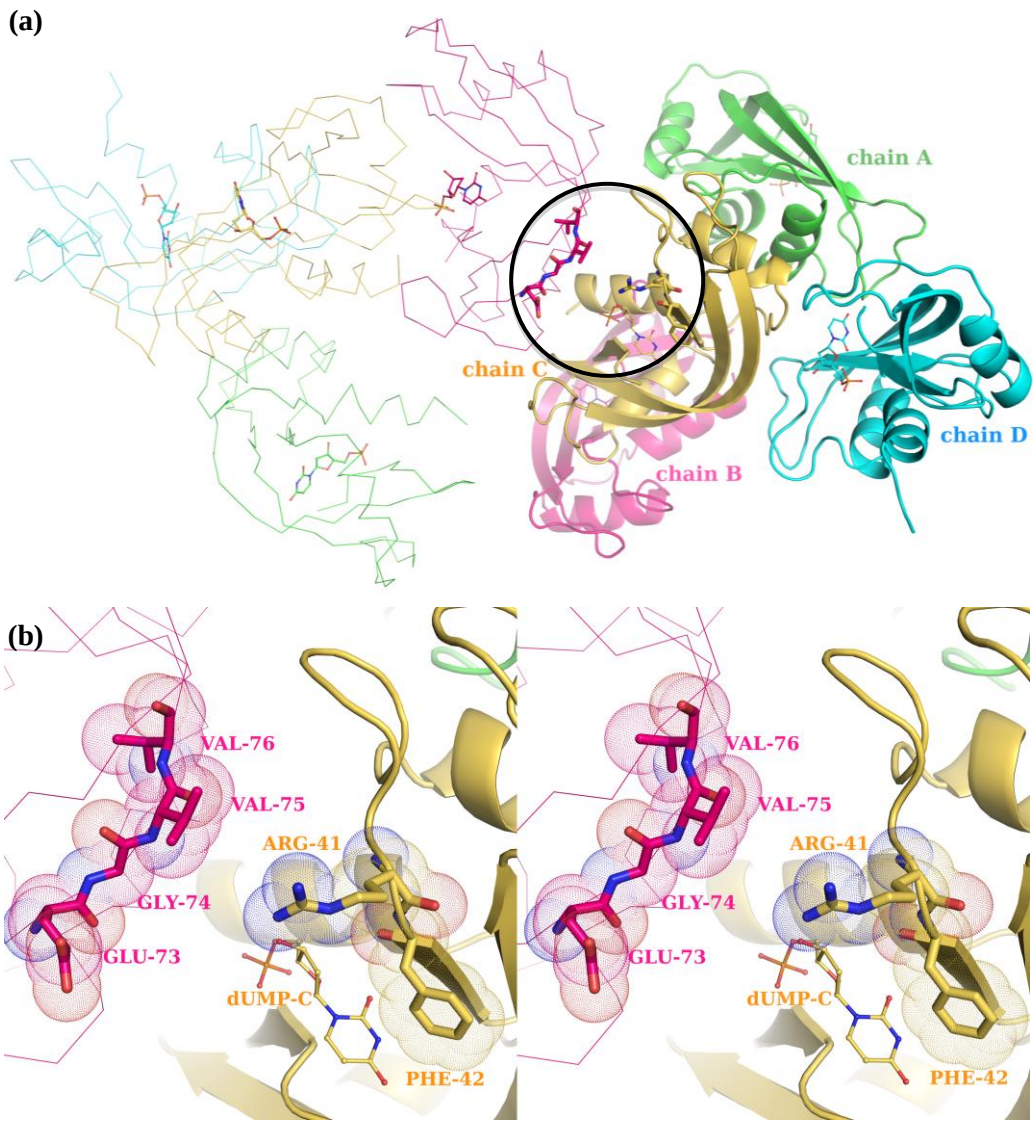


Figure 3.58. The conformation of Phe-42 in chain C of the PL_3 dUMP complex is not affected by crystal packing. (a) The four molecules of the PL_3 dUMP complex in the ASU showing in cartoon form with one of the adjacent ASU (showing in ribbon form), whose chain B is at the interface with chain C of the original ASU; (b) Close view of the interface near Arg-41 and Phe-42 in chain C of the PL_3 dUMP complex (yellow) and chain B in the adjacent ASU (pink). Glu-73, Gly-74, Val-75, and Val-76 in chain B are near the interface; however, they are too distanced apart to interact with Arg-41 or Phe-42 in chain C. Nitrogen and oxygen atoms are coloured in blue and red respectively; carbon atoms of chain B and chain C of the PL_3 dUMP complex are coloured in pink and yellow respectively.

Comparisons between PL_3 dUMP and RNase A U5P complexes

The cytosine preference of RNase A had led researchers to explore the possibility of using uridine-containing small molecules as RNase A inhibitors. Many crystal structures of RNase A complexed with natural and synthetic nucleotides have been solved, including the RNase A-U5P (uridine 5'-phosphate) structure (PDB: 3DXG; Tsirkone *et al.*, 2009). U5P and dUMP have a minimal conformational difference (the presence/absence of –OH group at the C2' position), therefore the structure of RNase A-U5P complex will be used in the comparison.

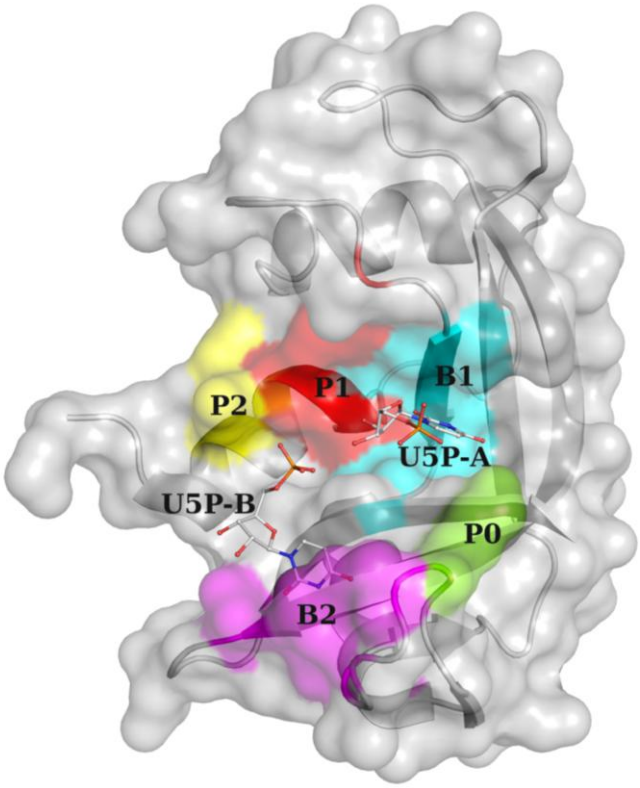


Figure 3.59. The three-dimensional structure of RNase A-U5P complex (PDB: 3DXG; Tsirkone *et al.*, 2009). Subsites P_0 (Lys-66), P_1 (Gln-11, His-12, Lys-41, His-119), P_2 (Lys-7, Arg-10), B_1 (Val-43, Asn-44, Thr-45, Phe-120, Ser-123), and B_2 (Asn-67, Gln-69, Asn-71, Glu-111) are labelled and marked with different colours. Two U5P ligands are bound to the RNase A molecule.

In the structure of RNase A-U5P complex, two U5P molecules (denoted U5P-A and U5P-B) bind to one RNase A molecule (Figure 3.59, 3.60). Only three potential hydrogen bonds are identified directly between U5P-A and His-12 and

Thr-45 of RNase A (Figure 3.60a), the corresponding residues in PL_3 (His-12 and Thr-44) are also involved in the interactions with dUMP ligand. However, more hydrogen bonds are observed between dUMP and PL_3 (Figure 3.49-3.50, Table 3.20-3.21) than between U5P and RNase A (Figure 3.60a). This is not a surprise, since U5P is a less favoured ligand of RNase A.

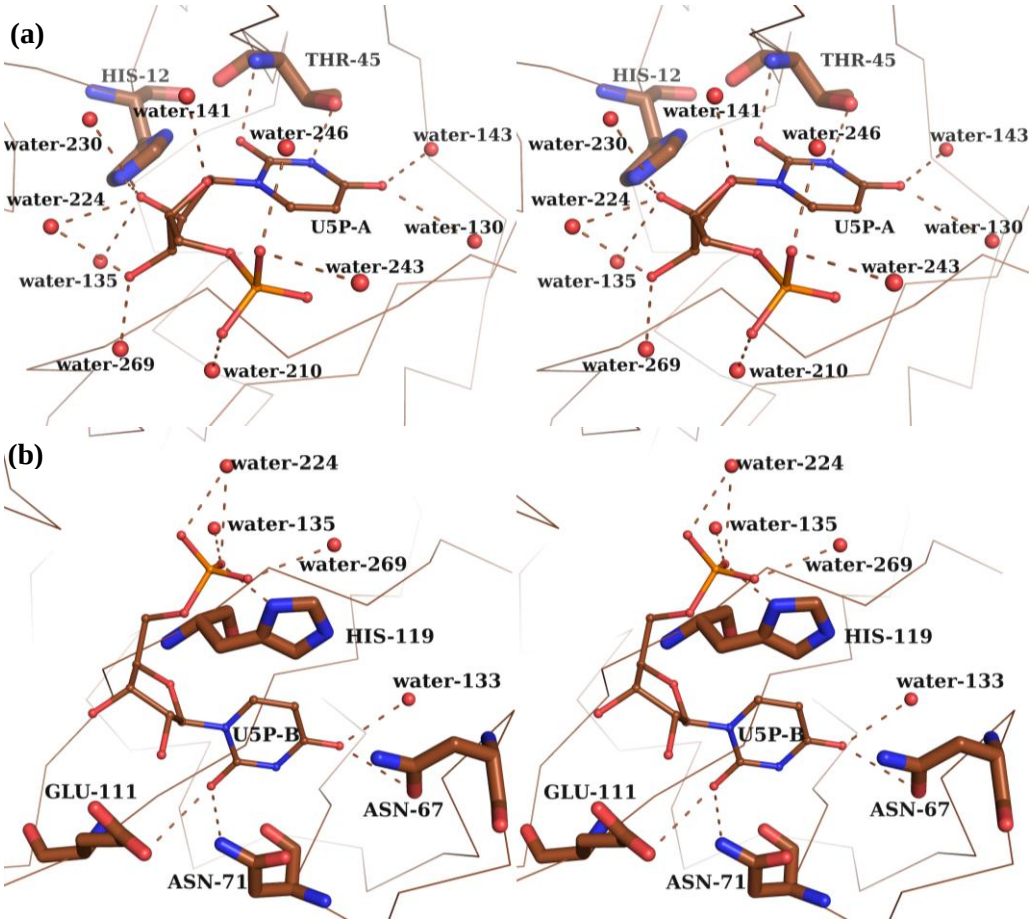


Figure 3.60. Interactions between (a) U5P-A and RNase A as well as (b) U5P-B and RNase A (PDB: 3DXG; Tsirkone *et al.*, 2009). Carbon, nitrogen and oxygen atoms are coloured in brown, blue and red respectively. The U5P ligands and RNase A residues interacting with the ligands are shown as sticks, water molecules are shown as spheres, potential hydrogen bonds are shown as dashed lines.

Unlike the various interactions observed between the phosphate groups of dUMP and PL_3, there are no direct interactions between RNase A and the phosphate group of U5P-A. This is due to the fact U5P-B occupying the P_1 and B_2 subsites of RNase A (Figure 3.60b) forces the phosphate group of U5P-A to move away from the protein, and stabilises U5P-A via water-mediated

interactions (Figure 3.60a). Such interactions between the ligand and the B_2 subsite are not observed in the PL_3 dUMP complex structure.

With further biological functions being discovered for RNase 4 (Cocchi *et al.*, 2012; Li *et al.*, 2013), it would be useful to study the interactions between this protein and cytosine-containing nucleotides kinetically and structurally, which would shed light on substrate recognition and possible inhibitor design for RNase 4.

3.4.3 Structure of PL_3 D80A dCMP complex

3.4.3.1 Overall structure of PL_3 D80A dCMP complex

The crystal structure of PL_3 D80A dCMP complex was determined at 2.6 Å resolution with four molecules in the ASU (Figure 3.61). The final model of the PL_3 D80A dCMP complex contains residues -1 – 119 in molecules A and B, residues 0 – 119 in molecules C and D, and 4 dCMP molecules. Ramachandran plot suggests that 95.1% of the residues in the structure of PL_3 D80A dCMP complex are in the most favourable regions with the remainder in the allowed regions (Figure 3.33 in Section 3.3.15.2).

One hydrogen bond is identified between molecules A-C and B-D respectively (Table 3.24). Analysis of the possibility of the formation of biological dimers were performed using the *PISA* web server (Krissinel and Henrick, 2007). A CSS score of 0 for all potential dimers in the ASU suggests that the observed arrangement of the PL_3 D80A dCMP molecules is not biological.

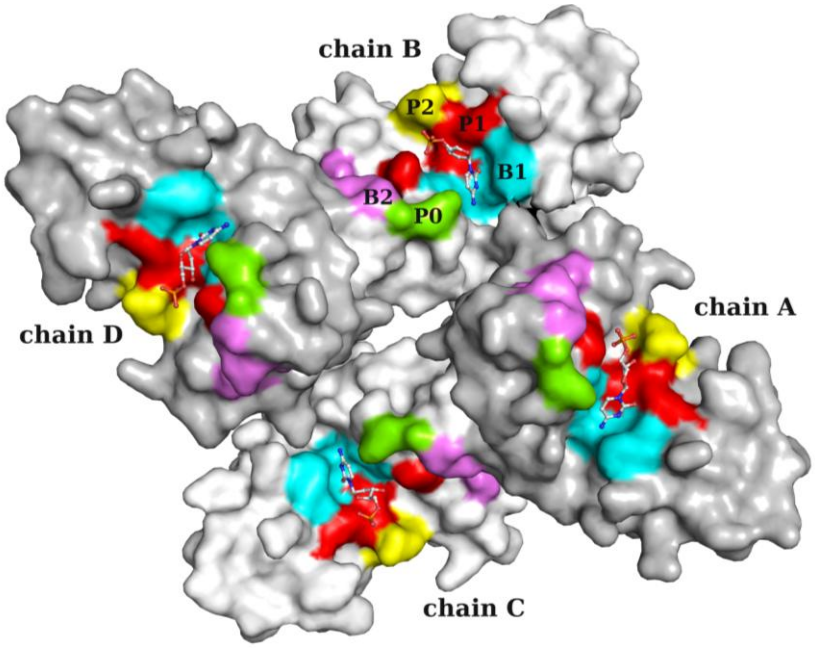


Figure 3.61. Arrangement of the four PL_3 D80A dCMP molecules in the ASU. dCMP ligands are shown in stick representation; The P_0, P_1, P_2, B_1, and B_2 subsites are coloured in green, red, yellow, cyan, and pink.

Table 3.24. Potential hydrogen bonds at various dimer interfaces within the ASU of the PL_3 D80A dCMP structure.

Hydrogen bond (donor – acceptor)	Distance (Å)
AVal-75 N – C Gly-67 O	3.23
D Val-75 N – B Gly-67 O	3.22

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

3.4.3.2 dCMP recognition by PL_3 D80A

In the structure of PL_3 D80A dCMP complex, the electron density maps of the dCMP nucleotides are well defined at the P_1 catalytic subsites of the PL_3 D80A proteins (Figure 3.62). Due to the initial high B-factors, the occupancies of the dCMP molecules were lowered (to 0.6 for dCMP-A and -B, 0.5 and 0.7 for the phosphate and nucleoside moieties of dCMP-C, and 0.3 for dCMP-D) so that they would have a similar B factor to those of the surrounding residues.

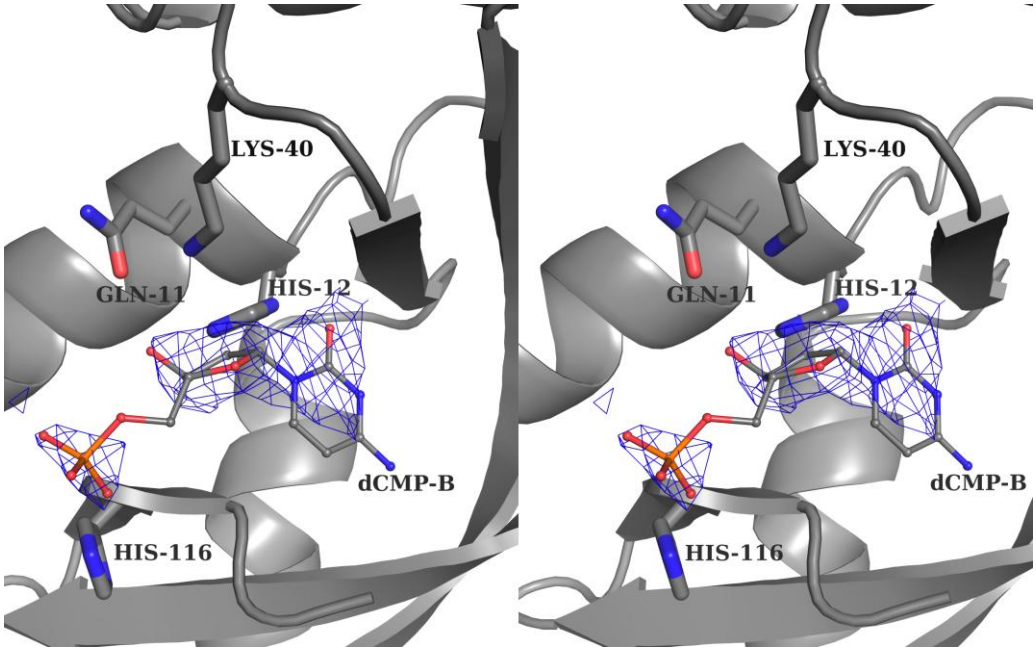


Figure 3.62. Close view of the P_1 catalytic subsite of PL_3 D80A dCMP complex (chain B) in stereo view. The dCMP ligand is shown as stick model with $2F_o$-F_c electron density map contoured at 1.0σ.

Three dCMP ligands (dCMP-B to -D) are in the preferred C3'-*endo* conformation, and dCMP-A is in the O4'-*endo* conformation (Moodie and Thornton, 1993) (examples of sugar puckers are shown in Figure 3.47). Similar to the interactions between dUMP and PL_3, the base moiety of the dCMP ligand locates in the B_1 subsite of PL_3 D80A, and the phosphate group is positioned near the P_1 subsite (Figure 3.63).

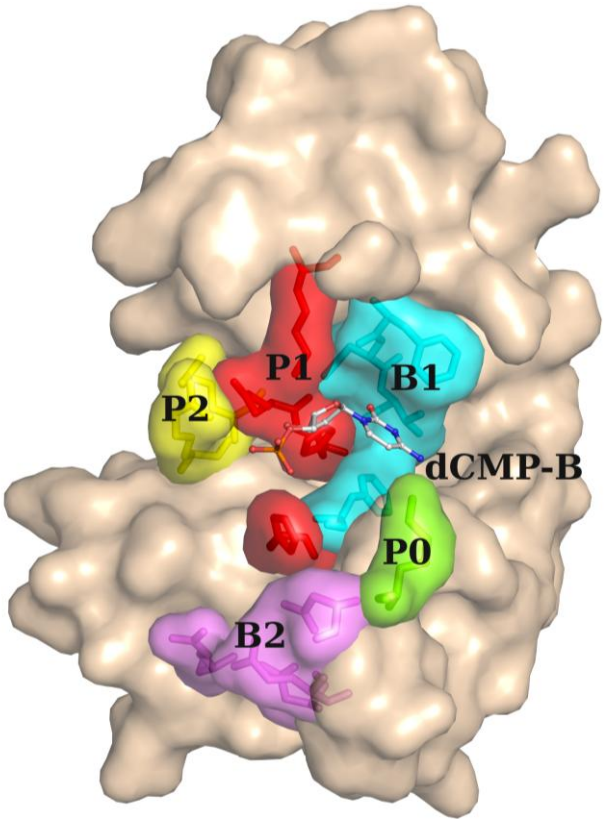


Figure 3.63. Surface diagram of the three-dimensional structure of PL_3 dUMP complex (chain B) with the subsites P_0 (Lys-65), P_1 (Gln-11, His-12, Lys-40, and His-116), P_2 (Arg-7 and Arg-10), B_1 (Phe-42, Asn-43, Thr-44, and Phe-117), and B_2 (Asn-66, Gln-68, Asn-70, and Glu-108) coloured in green, red, yellow, cyan, and pink respectively. Residues in the subsites as well as the dUMP-B ligand are shown in the stick representation. Carbon, nitrogen, oxygen, and phosphorus atoms of dUMP-B ligand are coloured in grey, blue, red and orange respectively.

The cytosine base of dCMP-A interacts with Lys-40 that belongs to the P_1 subsite (Figure 3.64, Table 3.25), whereas the base moiety of dCMP-B, -C and -D interacts with Thr-44 that is part of the B_1 subsite (Figure 3.65-3.67, Table 3.25-3.26). The only interaction between the phosphate group of dCMP and PL_3

D80A is observed in chain B: O1P of the phosphate group forms a hydrogen bond with His-116, one of the two catalytic His in the catalytic P_1 subsite (Figure 3.65, Table 3.25). In both chains C and D, residues in the P_1 subsite (His-12 of molecule C, and His-12 and Lys-40 of molecule D) interact with O2 of the ribose of dCMP ligands (Figure 3.66-3.67, Table 3.26). There are a few potential hydrogen bonds identified between the nitrogen atoms of the dCMP molecules and the protein molecules; however due to the unusual D-H…A angles limited by the resolution of the data, they are not listed here.

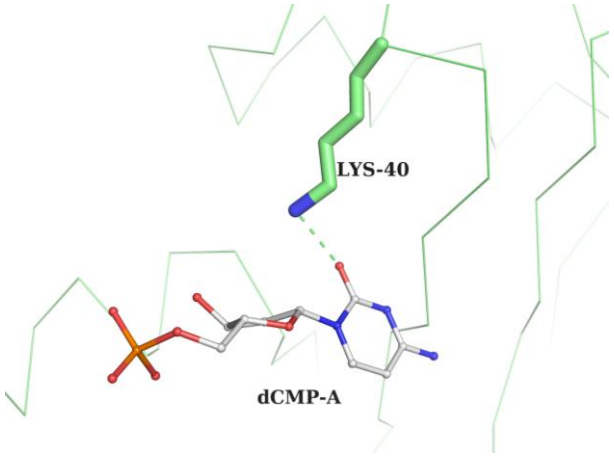


Figure 3.64. Details of the interactions at the catalytic site of PL_3 D80A molecule and dCMP in chain A. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain A of PL_3 D80A and dCMP-A are coloured in lime and grey respectively. Potential hydrogen bonds are shown as dashed lines.

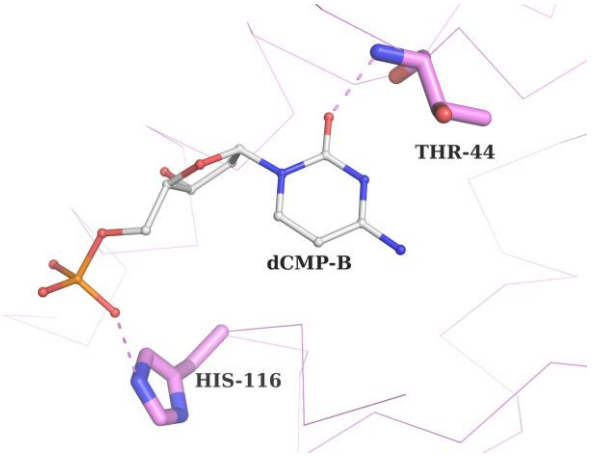


Figure 3.65. Details of the interactions at the catalytic site of PL_3 D80A molecule and dCMP in chain B. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain B of PL_3 D80A and dCMP-B are coloured in violet and grey respectively. Potential hydrogen bonds are shown as dashed lines.

Table 3.25. Potential hydrogen bonds between dCMP-1 and -2 ligands and their corresponding PL_3 D80A proteins.

dCMP-A atom	PL_3 D80A residue	Distance (Å)
O2	A Lys-40 NZ	2.61

dCMP-B atom	PL_3 D80A residue	Distance (Å)
O2	B Thr-44 N	2.53
O1P	B His-116 NE2	2.85

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

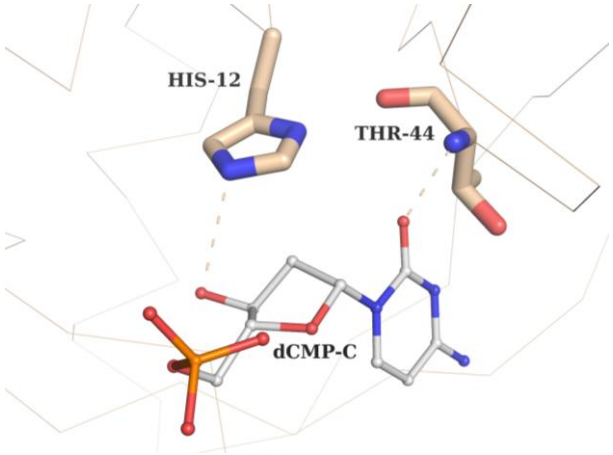


Figure 3.66. Details of the interactions at the catalytic site of PL_3 D80A molecule and dCMP in chain C. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain C of PL_3 D80A and dCMP-C are coloured in wheat and grey respectively. Potential hydrogen bonds are shown as dashed lines.

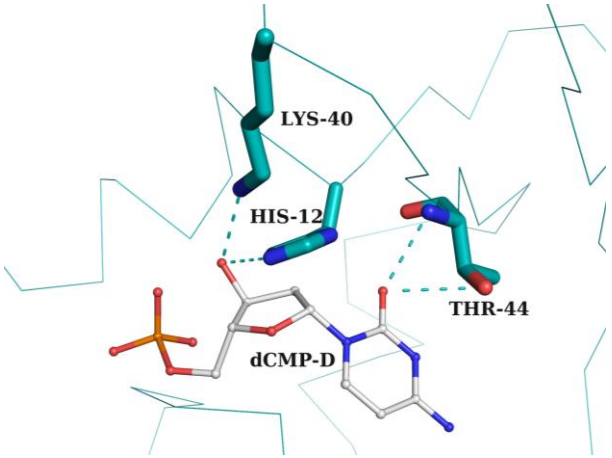


Figure 3.67. Details of the interactions at the catalytic site of PL_3 D80A molecule and dCMP in chain D. Nitrogen and oxygen atoms are coloured in blue and red; carbon atoms of the chain D of PL_3 D80A and dCMP-D are coloured in teal and grey respectively. Potential hydrogen bonds are shown as dashed lines.

Table 3.26. Potential hydrogen bonds between dCMP-3 and -4 ligands and their corresponding PL_3 D80A proteins.

dCMP-C atom	PL_3 D80A residue	Distance (Å)
O2	C Thr-44 N	2.35
O3'	C His-12 NE2	3.16

dCMP-D atom	PL_3 D80A residue	Distance (Å)
O2	D Thr-44 N	3.18
	D Thr-44 OG1	3.29
O3'	D His-12 NE2	3.21
	D Lys-40 NZ	2.57

Potential hydrogen bonds shorter than 3.3 Å with D-H…A angle > 120 °identified by *HBPLUS* (McDonald and Thornton, 1994) are shown.

3.4.3.3 Comparison of ligand recognition by PL_3 D80A with that by PL_3

Site directed mutagenesis studies of PL_3 reported by Hofsteenge *et al.* (1998a) revealed the importance of residue Asp-80 in the substrate specificity of the protein, and led us to begin the study of ligand recognitions by PL_3 and PL_3 D80A mutant from a structural point of view.

In the structure of PL_3 dUMP complex, the dUMP ligand is bound to the PL_3 protein by at least 6 hydrogen bonds. However, only 2 to 4 hydrogen bond interactions are observed between dCMP ligand and PL_3 D80A mutant. In all four dCMP molecules, the O2 atoms are all involved in hydrogen bonding, three of which with Thr-44, and the remainder with Lys-40.

In the structure of PL_3 dUMP complex, dUMP-B and -C have the highest occupancy, and PL_3 molecules B and C have lower B-factors therefore are more stable than the other two molecules. However, since the side chain conformation of Phe-42 of molecule B is affected by crystal packing (see Section 3.4.2.3), molecule C of the PL_3 dUMP complex structure will be referred to unless stated otherwise. Similarly, molecule C in the structure of PL_3 D80A dCMP complex will be referred to, due to its higher stability and ligand occupancy.

Comparison of the ligand-binding region of the structures of both PL_3 dUMP and PL_3 D80A dCMP complexes (with hRNase4 dUp complex structure as a reference) revealed some differences in the side chain conformations of the residues in or near the B_1 subsites (Figure 3.68). In the structures of PL_3 dUMP and hRNase4 dUp complexes (carbon atoms coloured in yellow and pink respectively), the side chain of Arg-101 is hydrogen bonded with the side chain of Asp-80, and held more towards the B_1 subsite, making Arg-101 more likely to interact with the =O group of uridine (as seen in the structure of the hRNase4 dUp complex). Also, in such conformation, the $–NH_2$ group of Arg-101 would strongly repulse the $–NH_2$ group of a cytosine base. In contrast, in the PL_3 D80A dCMP complex structure (carbon atoms coloured in cyan), the

attraction of the side chain of Arg-101 from Asp-80 is lost due to alanine substitution, and the side chain of Arg-101 is positioned further away from the B_1 subsite (Figure 3.68). This may reduce the repulsions that the $-NH_2$ groups of Arg-101 would apply to the $-NH_2$ group of the dCMP ligand, allowing the less favoured dCMP to occupy the B_1 subsite.

Another major difference observed between PL_3 and PL_3 D80A mutant is the χ_1 angle of His-116 (Figure 3.68). In the structure of PL_3 dUMP complex, the χ_1 angle of His-116 of the four molecules ranges from -56.8 °to -68.7 °, while in all the other three structures we present here (structures of PL_3, PL_3 D80A, and PL_3 D80A dCMP), the χ_1 angle of His-116 residues is between 158.8 ° and 176.2 °. The positive χ_1 angle of His-116 is also observed in the structures of free hRNase 4 and hRNase 4 dUp complex (Figure 3.68), as well as the majority of structures of RNase A in complex with a mononucleotide ligand. Such dual conformation of the catalytic His-119 (RNase A numbering) have been reported (Richards and Wykoff, 1971; de Mel *et al.*, 1994), and modelling studies have suggested that the two conformations His-119 of RNase A participate in the hydrolysis of RNA equally well (de Mel *et al.*, 1992).

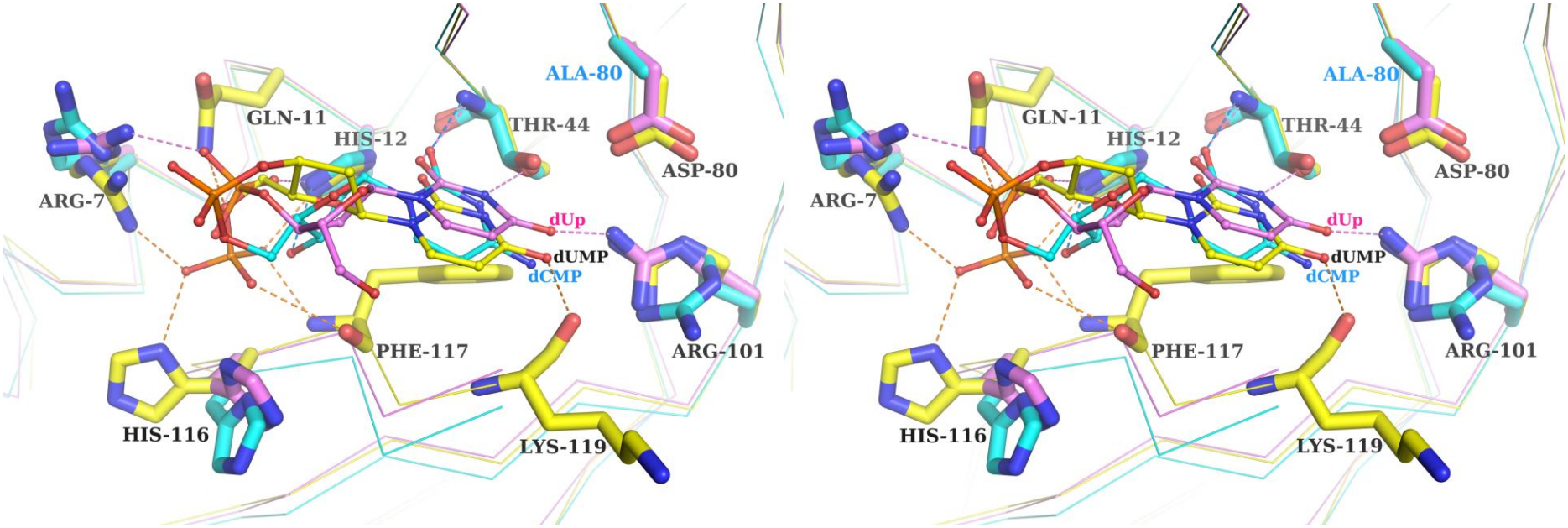


Figure 3.68. Stereo view of the ligand-binding regions of structures of PL_3 dUMP, PL_3 D80A dCMP, and hRNase 4 dUp complexes (PDB: 2RNF; Terzyan *et al.*, 1999). Nitrogen, oxygen, and phosphorus atoms are coloured in blue, red, and orange respectively; carbon atoms are coloured in yellow for PL_3 and dUMP, cyan for PL_3 D80A and dCMP, and pink for hRNase 4 and dUp. Potential hydrogen bonds are represented by dashed lines, and are coloured in orange for the PL_3 dUMP complex, blue for the PL_3 D80A dCMP complex, and pink for the hRNase 4 dUp complex.

3.5 Future Work

To date, we have successfully solved the structures of free porcine RNase 4 (PL_3), its D80A mutant, as well as two complex structures where PL_3 and PL_3 D80A bind their preferred mononucleotide ligand, dUMP and dCMP respectively. These structures have shed some light on the protein-ligand interactions of PL_3 and PL_3 D80A mutant; however, the change of substrate specificity caused by the D80A substitution cannot yet be fully explained. If more time was available, we would like to use two types of non-hydrolysable ligands (Table 3.27), which can be either dinucleotides or in the form of oligonucleotides, to form complexes with PL_3 and PL_3 D80A. In theory, the pyrimidine base and its 3' phosphate group should interact with the B_1 and P_1 subsites of PL_3 respectively and the adenosine binds to the B_2 subsite. These structures would allow us to further characterise the RNA recognition mechanism as well as the definitive role that Asp-80 plays in the substrate specificity of RNase 4.

Table 3.27. Non-hydrolysable ligands for substrate specificity studies of PL_3 and PL_3 D80A mutant.

Ligand binding to PL_3	Ligand binding to PL_3 D80A
2'-deoxyuridylyl-3', 5'-adenosine (2'-dUpA)	2'-deoxycytidylyl-3', 5'-adenosine (2'-dCpA)
2'-deoxy-2'-fluoro uridylyl-3', 5'-adenosine (2'-F-dUpA)	2'-deoxy-2'-fluoro cytidylyl-3', 5'-adenosine (2'-F-dCpA)

Another approach would be to use longer oligonucleotides containing one or two repeats of dUpA/dCpA as ligands, so that these complexes would mimic the recognition of large RNA molecules by PL_3 at some level. This would provide us with more knowledge on the interactions between the RNA and residues in different subsites of RNase, providing a better understanding of the mechanism of RNA recognition by this RNase molecule.

Chapter 4. Molecular recognition of porcine ribonuclease 4 (PL_3) by porcine ribonuclease inhibitor (pRI)

4.1 Ribonuclease inhibitor, natural inhibitor of RNases

Mammalian tissues express a potent proteinaceous inhibitor of the RNase A superfamily members. In 1952, Pirotte and Desreux first inferred the existence of such an inhibitor, from where started the extensive investigations on the nature (Roth, 1956; Roth, 1958a; Roth, 1958b; Roth, 1962; Shortman, 1961; Shortman, 1962) and distributions (Roth, 1962; Kraft and Shortman, 1970) of ribonuclease inhibitor (RI). Thanks to the novel method, which was based on the interactions between RNase and RI, developed in the late 1970s (Blackburn *et al.*, 1977), the isolation of quantities of homogeneous RI could be achieved, which in turn brought the research of RI into the modern era.

4.1.1 Primary structure of RI

RI is an acidic protein with a molecular weight of approximately 50 kDa characterised by a high content of leucine and cysteine (Blackburn *et al.*, 1977). To date, the primary structures of RI from pig (Hofsteenge *et al.*, 1988), human (Lee *et al.*, 1988; Schneider *et al.*, 1988), and rat (Kawanomoto *et al.*, 1992) have been determined by protein sequencing, and those of mouse, bovine, and chimpanzee RI were deduced from their complementary DNA (cDNA) sequences (Figure 4.1). These six proteins share a 60% sequence identity and an 86% sequence similarity. RIs from pig, cow, rat, and mouse have 456 amino acid residues, whereas human RI (hRI) and chimpanzee RI have 460 residues, both possessing a four-residue N-terminal extension.

Two interesting properties arose from the sequence analysis of RI. Firstly, RI contains an unusually large number of cysteine residues that are in sulphydryl form that makes these proteins highly sensitive to oxidation (Fominaya and Hofsteenge, 1992). Secondly, RI is constructed almost entirely of leucine-rich repeats (LRRs). LRRs are sequences of 24-29 amino acids that contain leucine or another hydrophobic amino acid at constant positions. Such repeats have been identified in more than 100 proteins with a diverse range of functions,

including DNA repair, RNA processing, cellular growth and development, enzyme inhibition or activation, signal transduction, extracellular matrix interactions, control of neural circuit formation, and bacterial virulence (Kobe and Deisenhofer, 1994; Buchanan and Gay, 1996; Hothorn *et al.*, 2011; She *et al.*, 2011; de Wit and Ghosh, 2014).

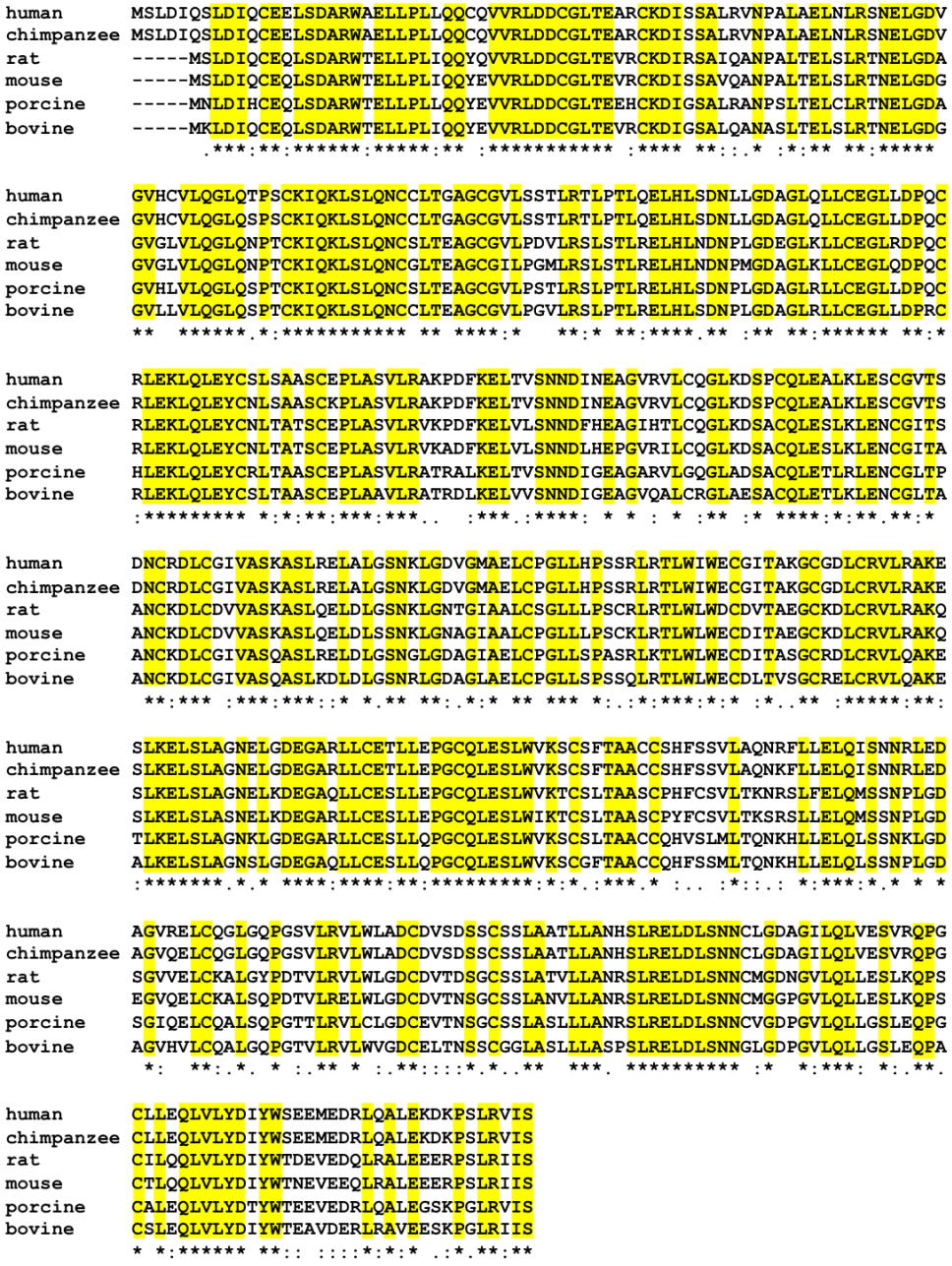


Figure 4.1. Comparison of the primary structures of human, chimpanzee, porcine, bovine, mouse, and rat RI. All sequences were taken from UniProtKB database. The multiple sequence alignment with a conservation scoring was performed using *Clustal Omega* EMBL-EBI online server (McWilliam *et al.*, 2013; Sievers *et al.*, 2011). Conserved residues are highlighted yellow.

4.1.2 Three-dimensional structure of RI and RI–RNase complexes

To date, crystal structures have been determined for free porcine RI (pRI) (Kobe and Deisenhofer, 1993) (Figure 4.2) and a few RI–RNase complexes, including pRI with RNase A (Kobe and Deisenhofer, 1995; Kobe and Deisenhofer, 1996), hRI with hRNase 1 (Johnson *et al.*, 2007), EDN (Iyer *et al.*, 2005), and ANG (Papageorgiou *et al.*, 1997), bovine RI with RNase A, chicken RI with chicken RNase 1, and mouse RI with mouse RNase 1 (Lomax *et al.*, 2014) (Figure 4.3).

pRI was the first LRR protein whose structure was ever determined. It revealed a novel non-globular supercoil in which 16 LRR units are symmetrically arranged into a horseshoe-like shape. The molecule has an overall dimension of 70 × 62 × 32 Å, and an inner diameter of 21 Å. Each LRR unit contains a short β strand and a longer α helix; the β strands form a parallel β sheet in the inner space of the horseshoe whereas the α helices are exposed to the environment. The conserved residues in the LRRs are important for the structure of RI. The leucine residues in each unit form a hydrophobic core between the β strand and the α helix, and between the LRR units. The side chains of asparagine and cysteine residues at position 10 of the repeat are buried and form hydrogen bonds to the polypeptide backbone.

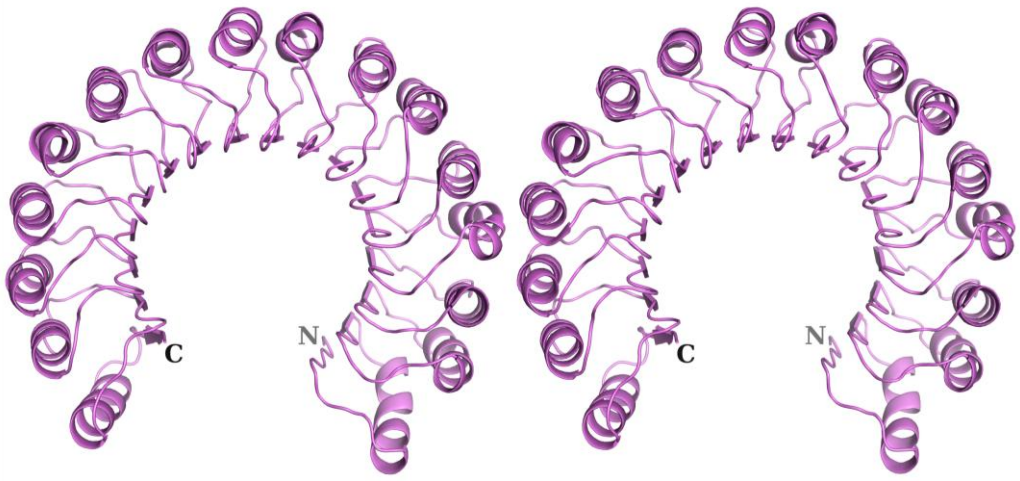


Figure 4.2. Stereo view of the crystal structure of pRI (PDB: 2BNH; Kobe and Deisenhofer, 1996). The N- and C- termini of pRI are labelled as N and C respectively.

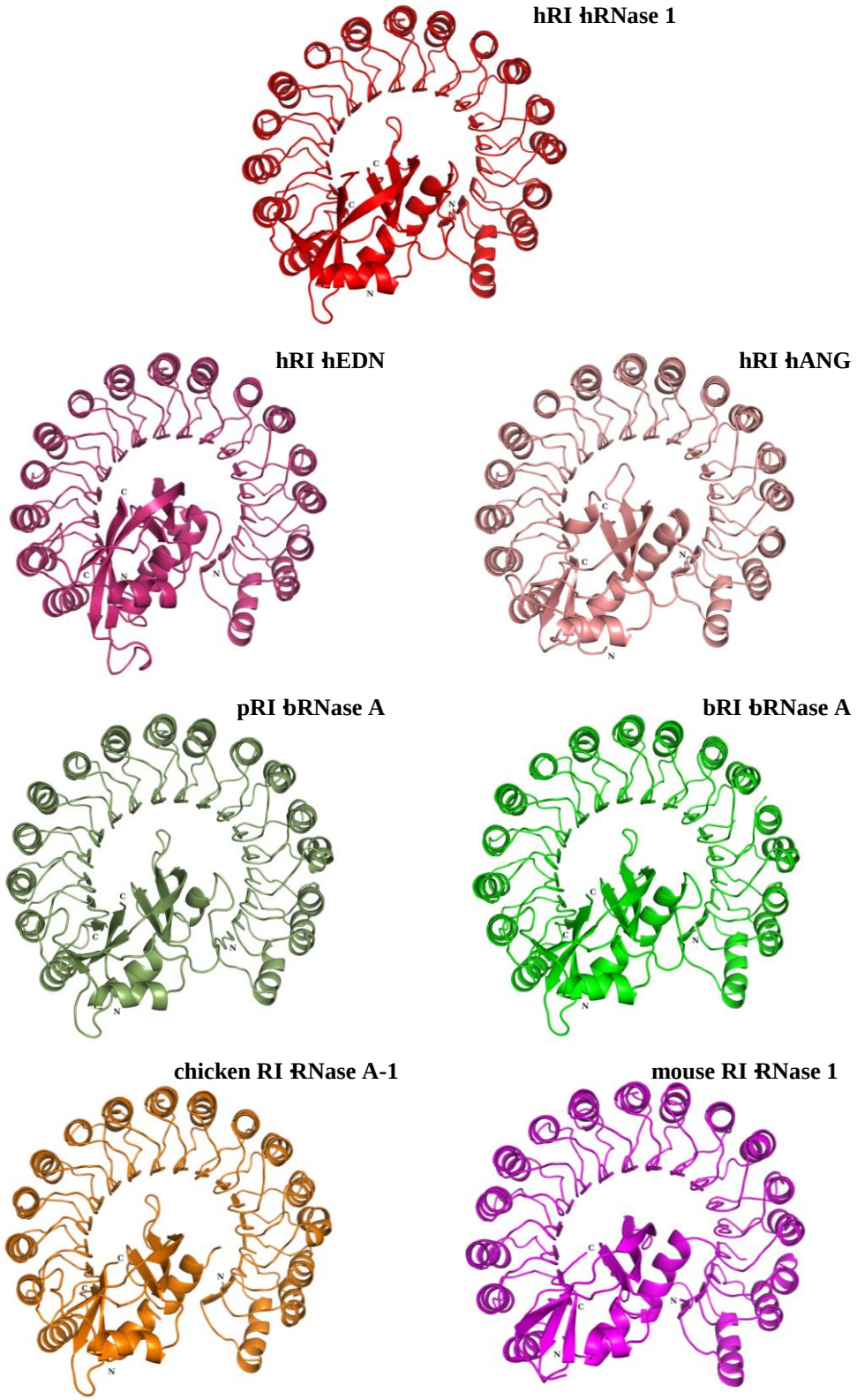
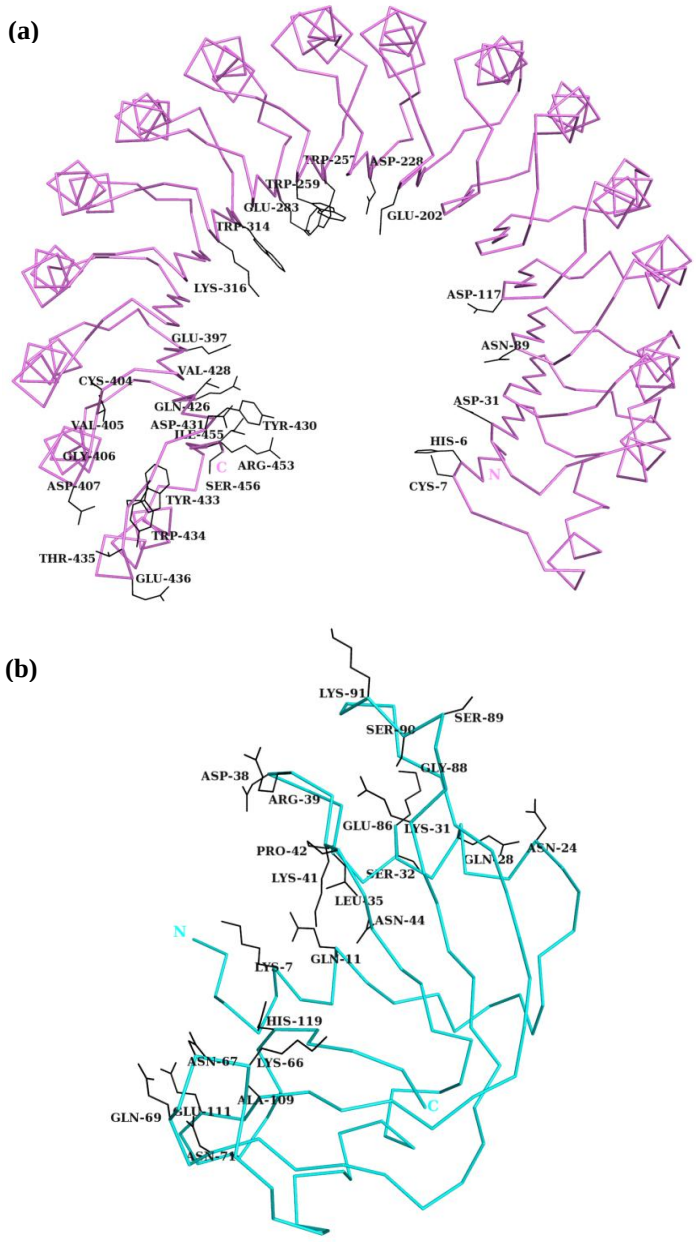


Figure 4.3. Structural similarity of RNase ƦRI complexes. hRI ɦRNase 1 (PDB: 1Z7X; Johnson *et al.*, 2007), hRI ɦEDN (PDB: 2BEX; Iyer *et al.*, 2005), hRI ɦANG (PDB: 1A4Y; Papageorgiou *et al.*, 1997), pRI ƦRNase A (PDB: 1DFJ; Kobe and Deisenhofer, 1995), bRI ƦRNase A, chicken RI ƦRNase A-1, and mouse RI ƦRNase 1 (PDB: 4PEQ, 4PER, 3TSR respectively; Lomax *et al.*, 2014) are shown.

RI inhibits RNase by blocking the access of the substrate to the active site (Kobe and Deisenhofer, 1995), and such inhibition is observed in all determined RI–RNase complex structures (Figure 4.3). The protein occupies the central cavity of the inhibitor contacting its C-terminal region more closely. In all seven complexes, the interface is large, involving at least 26 residues of RI that cluster in three separate groups and 24-30 residues from the protein ligand, an example for both components of the RI–RNase complex are shown in Figure 4.4.



(a) Residues on pRI that are involved in the recognition of RNase A. The side chains of pRI residues are shown as sticks and coloured in black; the N- and C-termini are labelled as N and C in violet;

(b) RNase A residues that bind to pRI. The side chains of RNase A residues are shown as sticks and coloured in black; the N- and C-termini are labelled as N and C in cyan.

Figure 4.4. C_α backbone of (a) pRI and (b) RNase A with the side chains of the contact residues shown in black. Figures were generated using the complex structure of pRI–RNase A (PDB: 1DFJ; Kobe and Deisenhofer, 1995).

Structural studies revealed that the C-terminal region of RI (residues 433 to 460; hRI numbering) is involved in the interactions with the catalytic site residues of RNases (Gln-11, Lys-41, and His-119, RNase A numbering); and the other contacts are required to stabilise the complexes. Site-specific mutagenesis of this region showed that residues Tyr-434, Asp-435, Tyr-437, and Ser-460 constitute a "hot spot" of binding energy (Chen and Shapiro, 1997; Chen and Shapiro, 1999).

4.1.3 Rationale behind this study

Preliminary analysis of pRNase 4 (PL_3) recognition by human RI revealed a 17-fold tighter binding than that of RNase A (Zelenko *et al.*, 1994). Despite the fact that residues Gln-11, Lys-41, and His-119 (RNase A numbering), which contribute the major portion of the binding energy in the interactions with RI, are also present in PL_3, it is difficult to provide an explanation for the tighter binding in the absence of the structure of PL_3 RI complex.

4.2 Modelling studies on PL_3 recognition by pRI

Modelling studies were carried out to gain some knowledge about the protein-protein interactions between PL_3 and pRI.

4.2.1 Methodology

Since the crystal structures of PL_3 and pRI are both available, a model of the PL_3 pRI complex structure was manually built in *PyMol*. Energy minimisation of the model was subsequently completed using the *YASARA* Energy Minimization Server (Krieger *et al.*, 2009). Energy minimisation is a necessary step in computational modelling, as it corrects the geometrical relationships between the two proteins, allows the side chains to place themselves in more comfortable positions, thus producing a model with the most favourable energy state.

4.2.2 Results

In the model of PL_3 pRI complex structure, PL_3 occupies the central cavity of pRI, making closer contact with the C-terminal region of the inhibitor (Figure 4.5), similar to the other RNase A superfamily members.

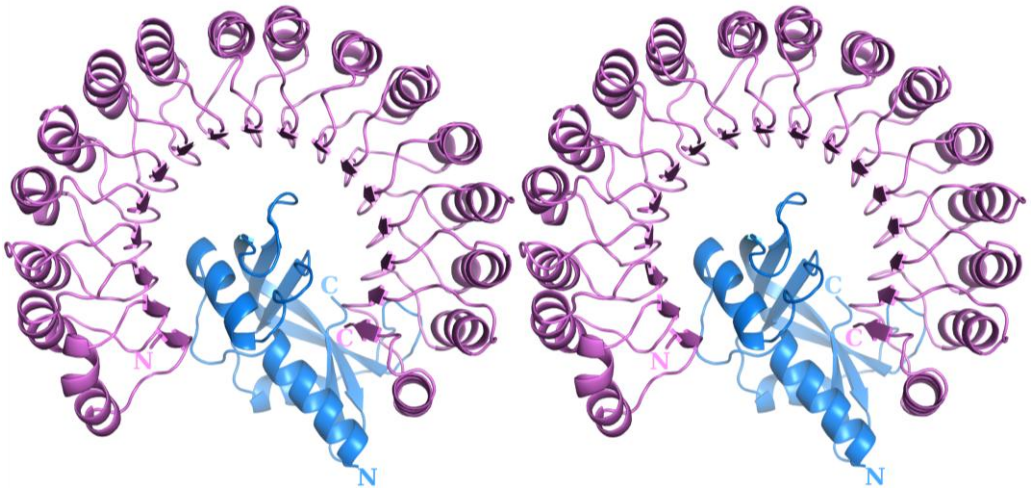


Figure 4.5. Stereo view of the energy minimised model of PL_3 pRI complex. The N- and C-termini of PL_3 (marine) and pRI (violet) are labelled as N and C respectively.

A simple analysis of the RNase A pRI complex with the model generated for the PL_3 pRI complex showed subtle differences in the orientation of side chains from the two protein complexes as shown in Figure 4.4 and Figure 4.6. However, a detailed three-dimensional structure of the complex is required to understand the molecular recognition of pRI by PL_3 molecule.

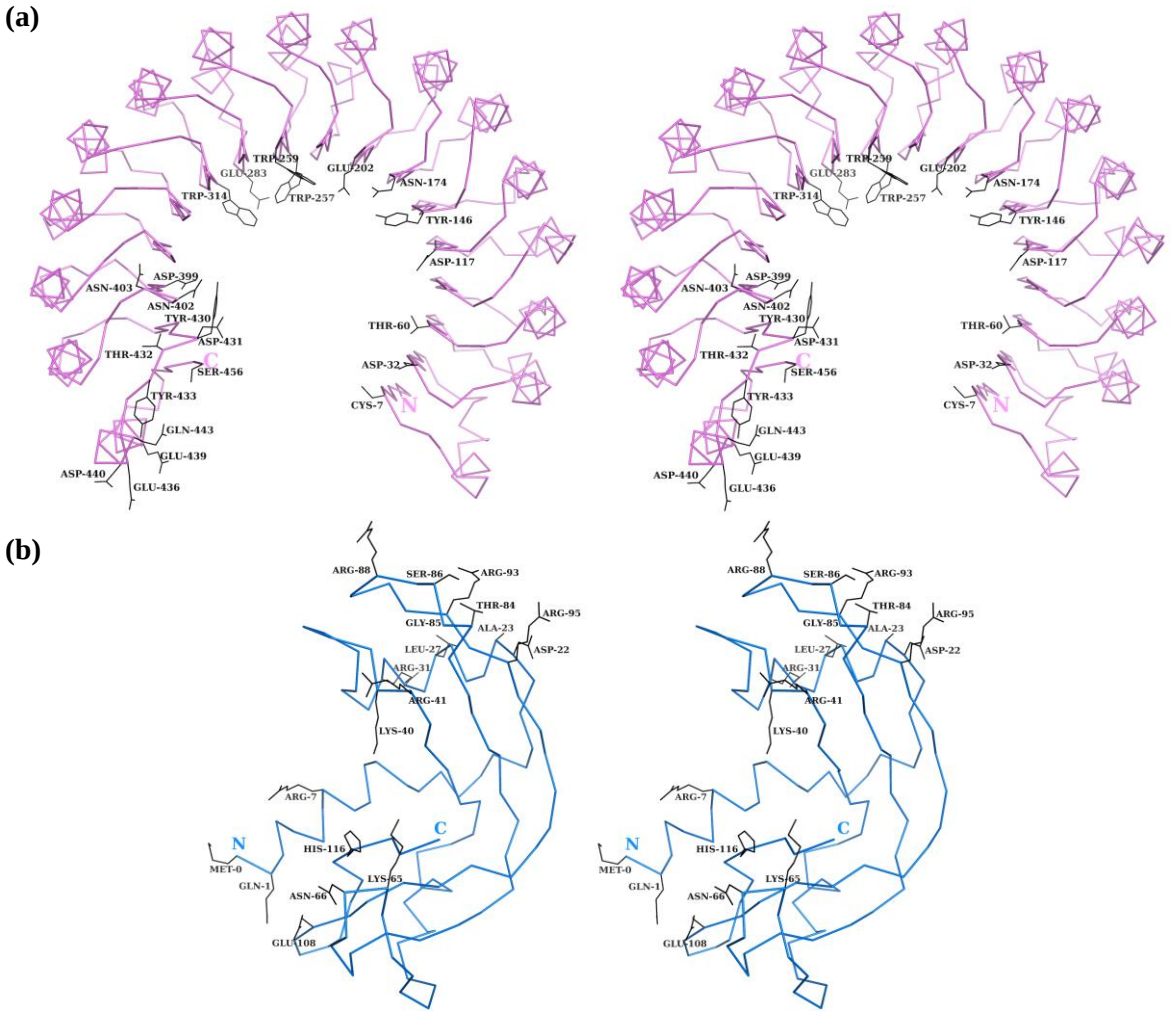


Figure 4.6. C_α backbone of (a) pRI and (b) PL_3 with the side chains of the contact residues shown in black. (a) Residues on pRI that are involved in the recognition of PL_3. The side chains of pRI residues are shown as sticks and coloured in black; the N- and C-termini are labelled as N and C in violet; (b) PL_3 Residues that bind to pRI. The side chains of PL_3 residues are shown as sticks and coloured in black; the N- and C-termini are labelled as N and C in marine.

4.3 Towards the crystal structure of PL_3 pRI complex

To validate the computer model, the structural characterisation of PL_3 pRI complex was pursued. Wild type PL_3 and pRI were expressed and purified, and the purified proteins were used for the formation of the complex.

4.3.1 Methods

4.3.1.1 Expression and purification of wild type PL_3

The BL21C$^+$(DE3)-RIPL-pET-46 EK/LIC-PL_3 construct was used for the expression of wild type PL_3. Optimised expression and purification protocols (Appendix II) were followed.

4.3.1.2 Expression and purification of pRI

The gene of pRI was carried by the ptRP vector in BL21(DE3) *E. coli* cells, which was obtained from our collaborator Dr Jan Hofsteenge from Basel, Switzerland. The protocol described by Kazakou (2008) was followed for the expression and purification of pRI.

LB broth medium supplemented with 100 μg/ml ampicillin was inoculated with 1% overnight culture, and incubated at 37 ℃ with shaking at 200 rpm until the cells had reached their late stationary phase, indicated by OD_{600} = 2.0. The cells were harvested by centrifuging at 15, 000 g (Beckman JLA10.500 rotor) and used to inoculate 1 L M9 minimal medium (1.5 g Pipes, 56.4 g minimal salt, 0.5 ml Tween 20, 1 ml glycerol per 1 L M9 media). The culture was incubated at 37 ℃ with shaking (200 rpm) for 6 hours before harvesting by centrifugation at 15, 000 g (Beckman JLA10.500 rotor). The cell pellet from 0.5 L culture was monodispersed in 50 ml lysis buffer (20 mM Tris-HCl pH 7.5); cells were

ruptured by cell disruptor at 20, 000 psi. The lysate was centrifuged at 70, 000 g (Beckman JA 25.50 rotor) for 33 minutes (Kazakou, 2008).

A 1 ml HiTrap NHS-activated HP column (GE Healthcare) was used to purify native pRI. The column was coupled with bovine pancreas ribonuclease A (RNase A) following instructions given in the manual of the column. The supernatant of cell lysate was loaded onto the column, which had been equilibrated with equilibration buffer (50 mM K_3PO_4 pH 6.4, 10 mM DTT, and 1 mM EDTA). The column was re-equilibrated with the same equilibration buffer after protein loading, subsequently washed with wash buffer (50 mM K_3PO_4 pH 6.4, 10 mM DTT, 1 mM EDTA, and 1 M NaCl). pRI was eluted from the column with elution buffer (0.1 M Na acetate pH 5.0, 0.3 M NaCl, 10 mM DTT, and 1 mM EDTA). Fractions with elevated UV trace were collected and dialysed against 2 L dialysis buffer (20 mM Tris-HCl pH7.5) at 4 ℃ overnight. The dialysed sample was analysed by SDS-PAGE electrophoresis (Kazakou, 2008).

4.3.1.3 Complex formation of PL_3·pRI

Purified pRI and PL_3 were mixed in a 1:1.2 molar ratio in 20 mM Tris-HCl pH 7.5, 20 mM DTT. The reaction mixture was incubated overnight at room temperature (Papageorgiou *et al.*, 1997).

4.3.1.4 Purification of PL_3·pRI complex by ion exchange chromatography

The purification of PL_3·pRI complex was carried out by ion exchange chromatography with a 1 ml HiTrap Q Sepharose FF column (GE Healthcare). The PL_3·pRI sample prepared by overnight incubation was centrifuged at 70, 000 g (Beckman 25.50 rotor) to remove any precipitate prior to purification, and subsequently loaded onto the HiTrap Q sepharose FF column which had been equilibrated with equilibration buffer (20 mM Tris-HCl pH 7.5). The

column was re-equilibrated with the same equilibration buffer, and the complex was eluted with an ascending NaCl gradient (0 to 0.4 M) in the equilibration buffer. Fractions with significant UV absorbance were collected and analysed by SDS-PAGE electrophoresis (Johnson *et al.*, 2007).

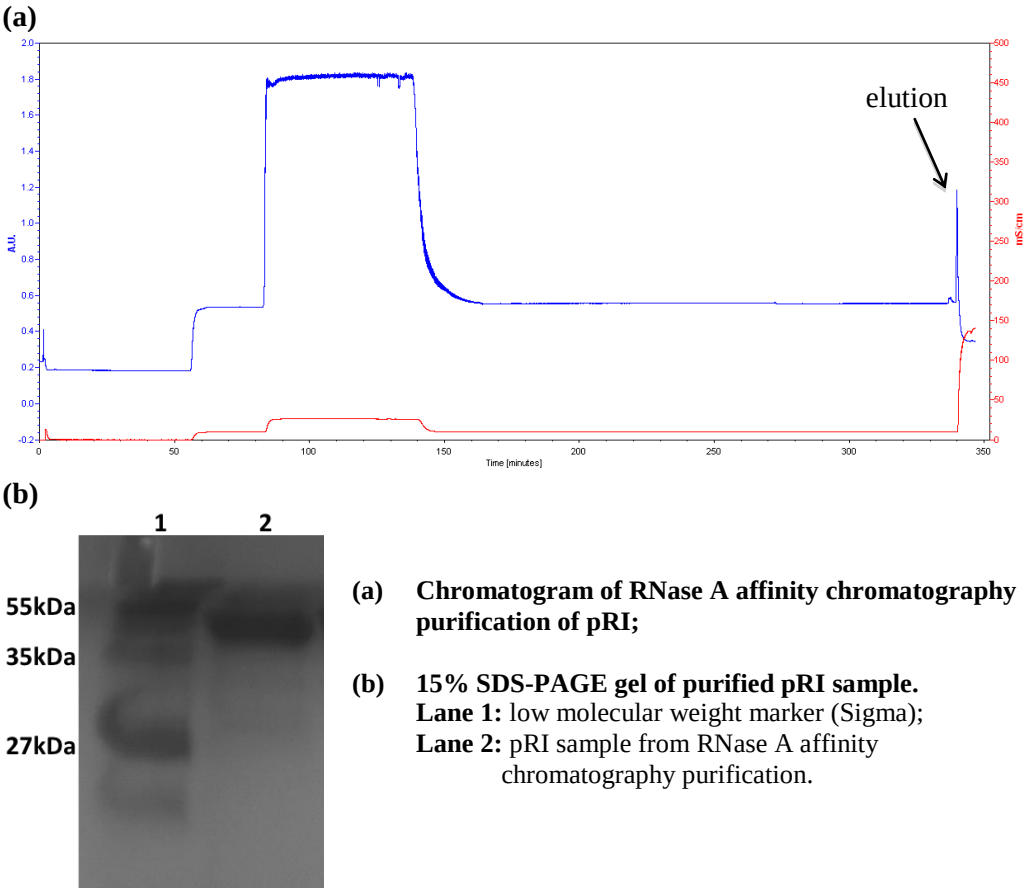
4.3.1.5 Purification of PL_3·pRI complex by size exclusion chromatography

An alternative method of purifying PL_3·pRI complex was carried out using size exclusion chromatography. The PL_3·pRI complex sample prepared by overnight incubation was centrifuged at 70, 000 g (Beckman 25.50 rotor) to remove any precipitate prior to purification, and subsequently purified using a Superdex 200 10/300GL size exclusion column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl. Fractions with significant UV absorbance were collected and analysed by SDS-PAGE electrophoresis (Kazakou, 2008).

4.3.2 Results and discussion

4.3.2.1 Expression and purification of pRI

The protocol described in Section 4.3.1 (Kazakou, 2008)(Kazakou, 2008)(Kazakou, 2008)(Kazakou, 2008)(Kazakou, 2008)(Kazakou, 2008) was followed for the expression and purification of pRI. The recombinant pRI protein was expressed in soluble form, and the RNase A affinity chromatography purification produced pure pRI (Figure 4.7). Approximately 3 mg pRI (at ~0.5 mg/ml) could be purified from 1 L bacterial culture.



(a) Chromatogram of RNase A affinity chromatography purification of pRI;

(b) 15% SDS-PAGE gel of purified pRI sample.
Lane 1: low molecular weight marker (Sigma);
Lane 2: pRI sample from RNase A affinity chromatography purification.

Figure 4.7. Results of pRI purification by RNase A affinity chromatography. Sample collected from the elution step in the (a) RNase A affinity chromatography purification was (b) electrophoresed under non-reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250.

4.3.2.2 Formation and purifications of PL_3·pRI complex

The protocol described in Section 4.3.1.3 was repeated twice to produce two samples that would be purified by anion exchange chromatography and size exclusion chromatography respectively. Unexpectedly, precipitation was observed in both samples. The precipitation was removed by centrifugation and the soluble fractions were collected.

One sample was purified by anion exchange chromatography as described in Section 4.3.1.4. Neither the chromatogram of the purification nor the SDS-PAGE gel showed sign of protein presence (Figure 4.8), suggesting the precipitants contained both pRI and PL_3.

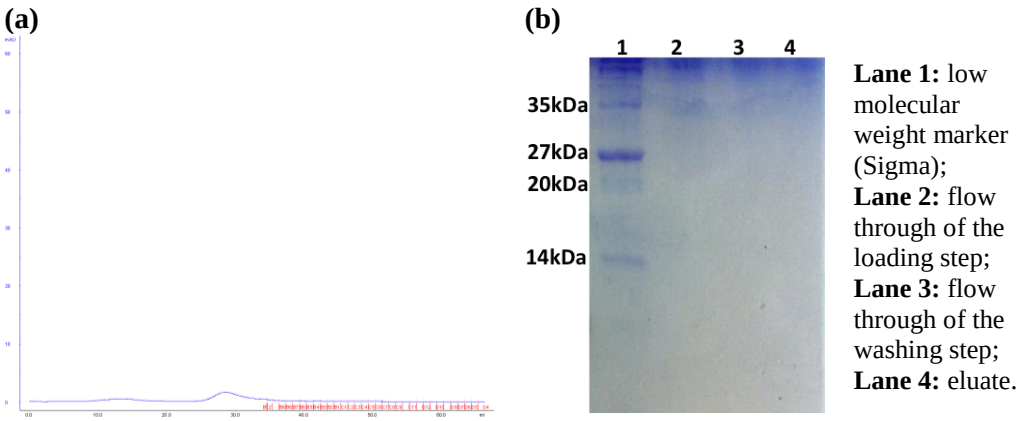


Figure 4.8. Results of Q-sepharose anion exchange purification of PL_3·pRI complex. (a) Q-sepharose anion exchange purification of PL_3·pRI; (b) 15% SDS-PAGE gel of fractions from Q-sepharose anion exchange purification.

Alternatively, the sample of PL_3·pRI complex was purified using size exclusion chromatography as described in Section 4.3.1.5. The chromatogram suggested that the sample consisted of three species, all of which were present at a low concentration (Figure 4.9a). The SDS-PAGE analysis of these three fractions however, suggested that no pRI was in the supernatant of the overnight-incubated sample (Figure 4.9b).

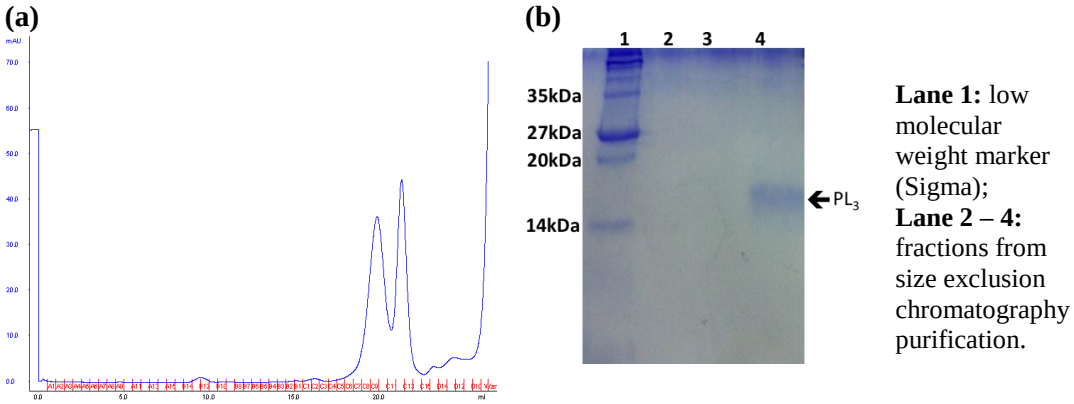


Figure 4.9. Results of size exclusion chromatography purification of PL_3·pRI complex. (a) Chromatogram of size exclusion chromatography purification of PL_3·pRI; (b) 15% SDS-PAGE gel of fractions from size exclusion chromatography purification. Protein samples were electrophoresed under reducing and denaturing conditions prior to staining with Coomassie brilliant blue R-250. Only tagged PL_3 protein was observed in the SDS-PAGE gel.

This part of the work was carried out using the wtPL_3 protein with the N-terminal His-tag still attached. The fact that both complex samples showed precipitations after incubation, together with previous observations (made during my Master's project) that the complex formation using a tag-free PL_3 purified from another construct did not result precipitation suggested that the N-terminal tag of PL_3 hindered the complex formation between PL_3 and pRI. This led to the His-tag removal experiments of PL_3. Due to lack of time, the complex formation using untagged PL_3 protein had not yet been carried out.

4.3.3 Future Work

The expression and purification protocols of pRI had been established; and quantities of pure pRI could be purified following these protocols. A new construct (BL21C^{+}(DE3)-RIPL-pET-46 Ek/LIC-PL_3) expressing PL_3 with cleavable N-terminal His-tag had been prepared in this study. Expression and purification protocols of PL_3 as well as the procedures for His-tag cleavage and subsequent purifications had been optimised; and pure untagged PL_3 could be produced with good yield. The next step would be use the untagged PL_3 protein to produce PL_3·pRI complex. Once the complex is formed and purified, crystallisation trials, structure determination and characterisation should follow, hoping to lead a better understanding of the interactions between RI and different RNases.

Chapter 5. General discussion

The RNase A superfamily members are a group of pyrimidine-specific ribonucleolytic enzymes with a wide distribution cross the animal kingdom. RNase 4 is an interesting member of this family, as it presents some distinguishing features from the others. RNase 4 is the shortest and the most conserved RNase, and it has a strong uridine preference compared with the rest of the family. However, RNase 4 is not as extensively studied as its homologues EDN, ECP, and ANG, which all show additional biological functions such as neurotoxicity and angiogenic activity.

Early kinetic and site-directed mutagenesis studies of porcine RNase 4 (also known as PL_3) suggested the importance of residues Asp-80 and Arg-101 in the substrate specificity of RNase 4 (Hofsteenge *et al.*, 1998a), which was supported by the crystal structures of hRNase 4 and its binary complex with dUp (Terzyan *et al.*, 1999). In the structure of hRNase 4 dUp complex, the side chain of Arg-101 is positioned away from the molecule surface, folded inwards and directed to the B_1 subsite, allowing hydrogen bonds to form with the ligand. Although Asp-80 does not form hydrogen bonds directly with the ligand, it was proposed that Asp-80 is forced to form a hydrogen bond with Thr-44 by the bulky side chain of Phe-42, and Thr-44 in turn interacts with the ligand.

We used PL_3 as a model to study the substrate recognition of RNase 4, and have solved four structures, namely PL_3, PL_3 D80A mutant, PL_3 dUMP complex, and PL_3 D80A dCMP complex.

Structural analyses show similarities in terms of protein-ligand interactions between PL_3 and hRNase 4: hydrogen bonds are identified between the ligand and residues Arg-7, His-12, Thr-44, and Phe-117 of both proteins. Significant differences are also observed, such as the absence of hydrogen bonds between Arg-101 of PL_3 and dUMP, and the interactions between dUMP and Gln-11, Lys-40, Asn-43, or Lys-119 of PL_3 that are not identified between dUp and hRNase 4. Although we observed stabilisation effect of Arg-101 side chain by

Asp-80, we could not yet fully address the level of involvement of residues Asp-80 and Arg-101 in substrate recognition due to the high flexibility as well as the different positions of phosphate groups of the ligands used. In order to get a more realistic picture of ligand recognition and substrate specificity of RNase 4, we would need to use oligonucleotide ligands to produce protein-ligand complex crystals.

Compared with the PL_3 dUMP complex, fewer hydrogen bonds are identified between the dCMP and PL_3 D80A mutant. However, considering the size and thus the flexibility of the dCMP ligand, it is likely that additional hydrogen bonds could be formed between the nitrogen atoms of dCMP and the PL_3 D80A protein. Since no structural information of PL_3 and PL_3 D80A mutant binding to the less favoured ligand (dCMP and dUMP respectively) is currently available, it is difficult to conclude whether the small number of hydrogen bonds between dCMP and D80A mutant is not enough for holding the ligand in the substrate-binding pocket.

In the PL_3 D80A dCMP complex structure, the side chain of Arg-101 is free to move away from the B_1 subsite in the absence of the side chain of Asp-80. It is likely that the increased distance between Arg-101 and the B_1 subsite reduces the repulsions between the $-NH_2$ groups of arginine and cytosine, allowing the B_1 subsite to host the less favoured (by the wild type) dCMP ligand. Thr-44 adopts the same conformation in PL_3 and PL_3 D80A mutant, suggesting Asp-80 is not crucial in holding Thr-44 in the observed position. However, the side chain of Asp-80 is involved in a hydrogen-bonding network in the B_1 subsite that is distorted in the D80A mutant, allowing the protein to bind the originally less-favoured cytidine ligand.

The activities of members in the RNase A superfamily can be abolished by the cytosolic ribonuclease inhibitor (RI), a 50 kDa protein consisting of leucine-rich repeats (LRRs) that arrange into a horseshoe-like shape. RI exists in many

animal species, and is involved in the regulation of cytoplasmic RNA level and protection of cells against unregulated RNase activity. RI binds to all the RNase A superfamily members, forming tight, reversible, one to one complexes. To date, several complex structures of RI-RNase from human, cow, chicken, and mouse have been determined, and all the structures revealed similar regions on the RNase molecules that are recognised by RI.

With the crystal structure of PL_3 being available, we could confirm the primary residues (Gln-11, Lys-41, and His-119 of RNase A) involved in RI-RNase interactions are also present in PL_3. However, the tighter binding of PL_3 to hRI than that of RNase A cannot be explained without a crystal structure of PL_3-RI complex.

Although we have not managed to produce the PL_3-RI protein complex, we determined that any additional amino acid residues at the N-terminal of PL_3, for example a fusion tag or a signal peptide, would hinder the complex formation. If more time was available, we should be able to produce and solve the structure of PL_3-RI complex, which would contribute to a better understanding of protein-protein interactions.

As D. C. Phillips and F. M. Richards stated in the preface of the book "Ribonuclease S", "no structure-analysis is ever complete" (Richards and Wyckoff, 1973). Every new structure solved presents researchers with interesting features and raises more questions. With the four structures we present here, the collection of RNase structures is more complete, and additional information on substrate recognition of RNase 4 is available. We would like to study the structures of PL_3 and PL_3 D80A mutant in complex with non-hydrolysable ligands, as well as the PL_3-RI protein complex, in order to further characterise the recognition of RNase 4 by its ligands as well as by RI, and provide some structural insight towards structure based drug design of inhibitors for different members of the RNase A superfamily.

References

Abtin, A., Eckhart, L., Mildner, M., Ghannadan, M., Harder, J., Schröder, J. M., Tschachler, E., 2009. Degradation by stratum corneum proteases prevents endogenous RNase inhibitor from blocking antimicrobial activities of RNase 5 and RNase 7. *J. Invest. Dermatol.* 129, 2193–201.

Acharya, K.R., Ackerman, S.J., 2014. Eosinophil granule proteins: form and function. *J. Biol. Chem.* 289, 17406–15.

Acharya, K.R., Shapiro, R., Allen, S.C., Riordan, J.F., Vallee, B.L., 1994. Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 91, 2915–9.

Acharya, K.R., Shapiro, R., Riordan, J.F., Vallee, B.L., 1995. Crystal structure of bovine angiogenin at 1.5 Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* 92, 2949–53.

Ackerman, S.J., Loegering, D.A., Venge, P., Olsson, I., Harley, J.B., Fauci, A.S., Gleich, G.J., 1983. Distinctive cationic proteins of the human eosinophil granule: major basic protein, eosinophil cationic protein, and eosinophil-derived neurotoxin. *J. Immunol.* 131, 2977–82.

Aguilar, C.F., Thomas, P.J., Mills, A., Moss, D.S., Palmer, R.A., 1992. Newly observed binding mode in pancreatic ribonuclease. *J. Mol. Biol.* 224, 265–7.

Anfinsen, C.B., 1973. Principles that govern the folding of protein chains. *Science* 181, 223–30.

Ardelt, W., Mikulski, S.M., Shogen, K., 1991. Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases. *J. Biol. Chem.* 266, 245–51.

Barnard, E.A., Stein, W.D., 1960. Some relations between enzyme activity, chemical reactivity and urea-induced disorientations in ribonuclease. *Biochim. Biophys. Acta* 37, 371–2.

Baynes, J.W., Wold, F., 1976. Effect of glycosylation on the in vivo circulating half-life of ribonuclease. *J. Biol. Chem.* 251, 6016–24.

Bedoya, V.I., Boasso, A., Hardy, A.W., Rybak, S., Shearer, G.M., Rugeles, M.T., 2006. Ribonucleases in HIV type 1 inhibition: effect of recombinant RNases on infection of primary T cells and immune activation-induced RNase gene and protein expression. *AIDS Res. Hum. Retroviruses* 22, 897–907.

Beintema, J.J., Hofsteenge, J., Iwama, M., Morita, T., Ohgi, K., Irie, M., Sugiyama, R.H., Schieven, G.L., Dekker, C.A., Glitz, D.G., 1988a. Amino acid sequence of the nonsecretory ribonuclease of human urine. *Biochemistry* 27, 4530–8.

Beintema, J.J., Schüller, C., Irie, M., Carsana, A., 1988b. Molecular evolution of the ribonuclease superfamily. *Prog. Biophys. Mol. Biol.* 51, 165–92.

Blackburn, P., Wilson, G., Moore, S., 1977. Ribonuclease inhibitor from human placenta. Purification and properties. *J. Biol. Chem.* 252, 5904–10.

Blake, C.C., Koenig, D.F., Mair, G.A., North, A.C., Phillips, D.C., Sarma, V.R., 1965. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* 206, 757–61.

Blow, D., 2002. *Outline of crystallography for biologists.* Oxford: Oxford University Press.

Blundell, T. L., Johnson, L. N., 1976. *Protein Crystallography.* London: Academic Press.

Boix, E., Leonidas, D.D., Nikolovski, Z., Nogués, M. V, Cuchillo, C.M., Acharya, K.R., 1999a. Crystal structure of eosinophil cationic protein at 2.4 Å resolution. *Biochemistry* 38, 16794–801.

Boix, E., Nikolovski, Z., Moiseyev, G.P., Rosenberg, H.F., Cuchillo, C.M., Nogués, M. V, 1999b. Kinetic and product distribution analysis of human eosinophil cationic protein indicates a subsite arrangement that favors exonuclease-type activity. *J. Biol. Chem.* 274, 15605–14.

Boix, E., Nogués, M.V., 2007. Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence. *Mol. Biosyst.* 3, 317–35.

Boix, E., Nogués, M. V, Schein, C.H., Benner, S.A., Cuchillo, C.M., 1994. Reverse transphosphorylation by ribonuclease A needs an intact p2-binding site. Point mutations at Lys-7 and Arg-10 alter the catalytic properties of the enzyme. *J. Biol. Chem.* 269, 2529–34.

Borah, B., Chen, C.W., Egan, W., Miller, M., Wlodawer, A., Cohen, J.S., 1985. Nuclear magnetic resonance and neutron diffraction studies of the complex of ribonuclease A with uridine vanadate, a transition-state analogue. *Biochemistry* 24, 2058–67.

Bragg, W.H., 1914. X-rays and crystalline structure. *Science* 40, 795–802.

Bragg, W.H., 1924. The analysis of crystal structure By X-Rays. *Science* 60, 139–49.

Brünger, A.T., 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–5.

Buchanan, S.G., Gay, N.J., 1996. Structural and functional diversity in the leucine-rich repeat family of proteins. *Prog. Biophys. Mol. Biol.* 65, 1–44.

Carreras, E., Boix, E., Navarro, S., Rosenberg, H.F., Cuchillo, C.M., Nogués, M.V., 2005. Surface-exposed amino acids of eosinophil cationic protein

play a critical role in the inhibition of mammalian cell proliferation. *Mol. Cell. Biochem.* 272, 1–7.

Carreras, E., Boix, E., Rosenberg, H.F., Cuchillo, C.M., Nogués, M.V., 2003. Both aromatic and cationic residues contribute to the membrane-lytic and bactericidal activity of eosinophil cationic protein. *Biochemistry* 42, 6636–44.

Castella, S., Fouchécourt, S., Teixeira-Gomes, A.P., Vinh, J., Belghazi, M., Dacheux, F., Dacheux, J.-L., 2004. Identification of a member of a new RNase a family specifically secreted by epididymal caput epithelium. *Biol. Reprod.* 70, 319–28.

Chen, C.Z., Shapiro, R., 1997. Site-specific mutagenesis reveals differences in the structural bases for tight binding of RNase inhibitor to angiogenin and RNase A. *Proc. Natl. Acad. Sci. U. S. A.* 94, 1761–6.

Chen, C.Z., Shapiro, R., 1999. Superadditive and subadditive effects of "hot spot" mutations within the interfaces of placental ribonuclease inhibitor with angiogenin and ribonuclease A. *Biochemistry* 38, 9273–85.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D. a, Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 12–21.

Cho, S., Beintema, J.J., Zhang, J., 2005. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* 85, 208–20.

Cho, S., Zhang, J., 2007. Zebrafish ribonucleases are bactericidal: implications for the origin of the vertebrate RNase A superfamily. *Mol. Biol. Evol.* 24, 1259–68.

Cocchi, F., DeVico, A.L., Lu, W., Popovic, M., Latinovic, O., Sajadi, M.M., Redfield, R.R., Lafferty, M.K., Galli, M., Garzino-Demo, A., Gallo, R.C.,

2012. Soluble factors from T cells inhibiting X4 strains of HIV are a mixture of β chemokines and RNases. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5411–6.

Cohen, S.N., Chang, A.C., Boyer, H.W., Helling, R.B., 1973. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 70, 3240–4.

Conforti, F.L., Sprovieri, T., Mazzei, R., Ungaro, C., La Bella, V., Tessitore, A., Patitucci, A., Magariello, A., Gabriele, A.L., Tedeschi, G., Simone, I.L., Majorana, G., Valentino, P., Condino, F., Bono, F., Monsurrò, M.R., Muglia, M., Quattrone, A., 2008. A novel Angiogenin gene mutation in a sporadic patient with amyotrophic lateral sclerosis from southern Italy. *Neuromuscul. Disord.* 18, 68–70.

Corrado, L., Battistini, S., Penco, S., Bergamaschi, L., Testa, L., Ricci, C., Giannini, F., Greco, G., Patrosso, M.C., Pileggi, S., Causarano, R., Mazzini, L., Momigliano-Richiardi, P., D'Alfonso, S., 2007. Variations in the coding and regulatory sequences of the angiogenin (ANG) gene are not associated to ALS (amyotrophic lateral sclerosis) in the Italian population. *J. Neurol. Sci.* 258, 123–7.

Crabtree, B., Holloway, D.E., Baker, M.D., Acharya, K.R., Subramanian, V., 2007. Biological and structural features of murine angiogenin-4, an angiogenic protein. *Biochemistry* 46, 2431–43.

Crick, F., 1970. Central dogma of molecular biology. *Nature* 227, 561–3.

Cuchillo, C.M., Parés, X., Guasch, a, Barman, T., Travers, F., Nogués, M. V, 1993. The role of 2',3'-cyclic phosphodiesters in the bovine pancreatic ribonuclease A catalysed cleavage of RNA: intermediates or products? *FEBS Lett.* 333, 207–10.

D'Alessio, G., Di Donato, A., Parente, A., Piccoli, R., 1991. Seminal RNase: a unique member of the ribonuclease superfamily. *Trends Biochem. Sci.* 16, 104–6.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., Richardson, D.C., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375–83.

DeLano, W. L., 2002. *The PyMOL Molecular Graphic System.* http://www.pymol.org.

De Mel, V.S., Doscher, M.S., Martin, P.D., Edwards, B.F., 1994. The occupancy of two distinct conformations by active-site histidine-119 in crystals of ribonuclease is modulated by pH. *FEBS Lett.* 349, 155–60.

De Mel, V.S., Martin, P.D., Doscher, M.S., Edwards, B.F., 1992. Structural changes that accompany the reduced catalytic efficiency of two semisynthetic ribonuclease analogs. *J. Biol. Chem.* 267, 247–56.

De Wit, J., Ghosh, A., 2014. Control of neural circuit formation by leucine-rich repeat proteins. *Trends Neurosci.* 1–12.

Deakyne, C.A., Allen, L.C., 1979. Role of active-site residues in the catalytic mechanism of ribonuclease A. *J. Am. Chem. Soc.* 101, 3951–3959.

delCardayré, S.B., Raines, R.T., 1994. Structural determinants of enzymatic processivity. *Biochemistry* 33, 6031–7.

delCardayré, S.B., Raines, R.T., 1995. A residue to residue hydrogen bond mediates the nucleotide specificity of ribonuclease A. *J. Mol. Biol.* 252, 328–36.

delCardayré, S.B., Ribó, M., Yokel, E.M., Quirk, D.J., Rutter, W.J., Raines, R.T., 1995. Engineering ribonuclease A: production, purification and characterization of wild-type enzyme and mutants at Gln11. *Protein Eng.* 8, 261–73.

Domachowske, J.B., Bonville, C.A., Dyer, K.D., Rosenberg, H.F., 1998a. Evolution of antiviral activity in the ribonuclease A gene superfamily:

evidence for a specific interaction between eosinophil-derived neurotoxin (EDN/RNase 2) and respiratory syncytial virus. *Nucleic Acids Res.* 26, 5327–32.

Domachowske, J.B., Dyer, K.D., Bonville, C.A., Rosenberg, H.F., 1998b. Recombinant human eosinophil-derived neurotoxin/RNase 2 functions as an effective antiviral agent against respiratory syncytial virus. *J. Infect. Dis.* 177, 1458–64.

Dubos, R.J., 1937. the Decomposition of Yeast Nucleic Acid By a Heat Resistant Enzyme. *Science* 85, 549–50.

Durack, D.T., Ackerman, S.J., Loegering, D.A., Gleich, G.J., 1981. Purification of human eosinophil-derived neurotoxin. *Proc. Natl. Acad. Sci. U. S. A.* 78, 5165–9.

Dyer, K.D., Rosenberg, H.F., Zhang, J., 2004. Isolation, characterization, and evolutionary divergence of mouse RNase 6: evidence for unusual evolution in rodents. *J. Mol. Evol.* 59, 657–65.

Eftink, M. R., Biltonen, R. L., 1987. Pancreatic ribonuclease A: The most studied endoribonuclease. In: A. Neuberger, K. Brocklehurst, eds. *Hydrolytic Enzymes.* Amsterdam: Elsevier. pp. 333-376.

Egesten, a, Dyer, K.D., Batten, D., Domachowske, J.B., Rosenberg, H.F., 1997. Ribonucleases and host defense: identification, localization and gene expression in adherent monocytes in vitro. *Biochim. Biophys. Acta* 1358, 255–60.

Eisenstein, M., 2006. A look back: adventures in the matrix. *Nat. Methods* 3, 410–410.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66, 486–501.

Evans, P., 2006. Scaling and assessment of data quality. *Acta Crystallogr. D. Biol. Crystallogr.* 62, 72–82.

Evans, P.R., 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D. Biol. Crystallogr.* 67, 282–92.

Fankuchen, I., 1941. An X-ray and crystallographic study of ribonuclease. *J. Gen. Physiol.* 24, 315–6.

Fernández-Santiago, R., Hoenig, S., Lichtner, P., Sperfeld, A.-D., Sharma, M., Berg, D., Weichenrieder, O., Illig, T., Eger, K., Meyer, T., Anneser, J., Münch, C., Zierz, S., Gasser, T., Ludolph, A., 2009. Identification of novel Angiogenin (ANG) gene missense variants in German patients with amyotrophic lateral sclerosis. *J. Neurol.* 256, 1337–42.

Fett, J.W., Strydom, D.J., Lobb, R.R., Alderman, E.M., Bethune, J.L., Riordan, J.F., Vallee, B.L., 1985. Isolation and characterization of angiogenin, an angiogenic protein from human carcinoma cells. *Biochemistry* 24, 5480–6.

Findlay, D., Herries, D.G., Mathias, A.P., Rabin, B.R., Ross, C.A., 1961. The active site and mechanism of action of bovine pancreatic ribonuclease. *Nature* 190, 781–84.

Fisher, B.M., Ha, J.H., Raines, R.T., 1998. Coulombic forces in protein-RNA interactions: binding and cleavage by ribonuclease A and variants at Lys7, Arg10, and Lys66. *Biochemistry* 37, 12121–32.

Flogel, M., Albert, a, Biltonen, R., 1975. The magnitude of electrostatic interactions in inhibitor binding and during catalysis by ribonuclease A. *Biochemistry* 14, 2616–21.

Fominaya, J.M., Hofsteenge, J., 1992. Inactivation of ribonuclease inhibitor by thiol-disulfide exchange. *J. Biol. Chem.* 267, 24655–60.

Fredens, K., Dahl, R., Venge, P., 1982. The Gordon phenomenon induced by the eosinophil cationic protein and eosinophil protein X. *J. Allergy Clin. Immunol.* 70, 361–6.

Fu, D., Chen, L., O'Neill, R. a, 1994. A detailed structural characterization of ribonuclease B oligosaccharides by 1H NMR spectroscopy and mass spectrometry. *Carbohydr. Res.* 261, 173–86.

Futami, J., Tsushima, Y., Murato, Y., Tada, H., Sasaki, J., Seno, M., Yamada, H., 1997. Tissue-specific expression of pancreatic-type RNases and RNase inhibitor in humans. *DNA Cell Biol.* 16, 413–9.

Garman, E., Owen, R.L., 2007. Cryocrystallography of macromolecules: practice and optimization. *Methods Mol. Biol.* 364, 1–18.

Garvey, T.L., Dyer, K.D., Ellis, J.A., Bonville, C.A., Foster, B., Prussin, C., Easton, A.J., Domachowske, J.B., Rosenberg, H.F., 2005. Inflammatory responses to pneumovirus infection in IFN-alpha beta R gene-deleted mice. *J. Immunol.* 175, 4735–44.

Gegenheimer, P., Apirion, D., 1981. Processing of procaryotic ribonucleic acid. *Microbiol. Rev.* 45, 502–41.

Gellera, C., Colombrita, C., Ticozzi, N., Castellotti, B., Bragato, C., Ratti, A., Taroni, F., Silani, V., 2008. Identification of new ANG gene mutations in a large cohort of Italian patients with amyotrophic lateral sclerosis. *Neurogenetics* 9, 33–40.

Gleich, G.J., Loegering, D.A., Bell, M.P., Checkel, J.L., Ackerman, S.J., McKean, D.J., 1986. Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 83, 3146–50.

Green, D.W., Ingram, V.M., Perutz, M.F., 1954. The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 225, 287–307.

Greenway, M.J., Alexander, M.D., Ennis, S., Traynor, B.J., Corr, B., Frost, E., Green, A., Hardiman, O., 2004. A novel candidate region for ALS on chromosome 14q11.2. *Neurology* 63, 1936–8.

Greenway, M.J., Andersen, P.M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J., Morrison, K.E., Green, A., Acharya, K.R., Brown, R.H., Hardiman, O., 2006. ANG mutations segregate with familial and "sporadic" amyotrophic lateral sclerosis. *Nat. Genet.* 38, 411–3.

Gundlach, H.G., Stein, W.H., Moore, S., 1959. The nature of the amino acid residues involved in the inactivation of ribonuclease by iodoacetate. *J. Biol. Chem.* 234, 1754–60.

Gupta, V., Muyldermans, S., Wyns, L., Salunke, D.M., 1999. The crystal structure of recombinant rat pancreatic RNase A. *Proteins* 35, 1–12.

Hamann, K.J., Gleich, G.J., Checkel, J.L., Loegering, D.A., McCall, J.W., Barker, R.L., 1990. In vitro killing of microfilariae of Brugia pahangi and Brugia malayi by eosinophil granule proteins. *J. Immunol.* 144, 3166–73.

Hammes, G.G., 1968. Relaxation spectrometry of biological systems. *Adv. Protein Chem.* 23, 1–57.

Harder, J., Schroder, J.-M., 2002. RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin. *J. Biol. Chem.* 277, 46779–84.

Hendrickson, W., 1991. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* 254, 51–58.

Hirs, C.H., 1962. Dinitrophenyiribonucleases. *Brookhaven Symp. Biol.* 15, 154–83.

Hofsteenge, J., 1994. "Holy" proteins. I: Ribonuclease inhibitor. *Curr. Opin. Struct. Biol.* 4, 807–9.

Hofsteenge, J., Kieffer, B., Matthies, R., Hemmings, B.A., Stone, S.R., 1988. Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals the presence of leucine-rich repeats. *Biochemistry* 27, 8537–44.

Hofsteenge, J., Matthies, R., Stone, S.R., 1989. Primary structure of a ribonuclease from porcine liver, a new member of the ribonuclease superfamily. *Biochemistry* 28, 9806–13.

Hofsteenge, J., Moldow, C., Vicentini, A.M., Zelenko, O., Jarai-Kote, Z., Neumann, U., 1998a. A single amino acid substitution changes ribonuclease 4 from a uridine-specific to a cytidine-specific enzyme. *Biochemistry* 37, 9250–7.

Hofsteenge, J., Vicentini, A., Zelenko, O., 1998b. Ribonuclease 4, an evolutionarily highly conserved member of the superfamily. *Cell. Mol. Life Sci.* 54, 804–10.

Holloway, D.E., Chavali, G.B., Hares, M.C., Baker, M.D., Subbarao, G. V, Shapiro, R., Acharya, K.R., 2004. Crystallographic studies on structural features that determine the enzymatic specificity and potency of human angiogenin: Thr44, Thr80, and residues 38-41. *Biochemistry* 43, 1230–41.

Holloway, D.E., Chavali, G.B., Hares, M.C., Subramanian, V., Acharya, K.R., 2005. Structure of murine angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies. *Acta Crystallogr. D. Biol. Crystallogr.* 61, 1568–78.

Holm, L., Park, J., 2000. DaliLite workbench for protein structure comparison. *Bioinformatics* 16, 566–7.

Hooper, L. V, Stappenbeck, T.S., Hong, C. V, Gordon, J.I., 2003. Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat. Immunol.* 4, 269–73.

Hosoya, K., Nagareda, Y., Hasemi, S., Sanda, A., Takizawa, Y., Watanabe, H., Ohgi, K., Irie, M., 1990. Primary structure of an alkaline ribonuclease from bovine liver. *J. Biochem.* 107, 613–8.

Hothorn, M., Belkhadir, Y., Dreux, M., Dabi, T., Noel, J.P., Wilson, I.A., Chory, J., 2011. Structural basis of steroid hormone perception by the receptor kinase BRI1. *Nature* 474, 467–71.

Huang, Y.-C., Lin, Y.-M., Chang, T.-W., Wu, S.-J., Lee, Y.-S., Chang, M.D.-T., Chen, C., Wu, S.-H., Liao, Y.-D., 2007. The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. *J. Biol. Chem.* 282, 4626–33.

Ishihara, K., Asai, K., Nakajima, M., Mue, S., Ohuchi, K., 2003. Preparation of recombinant rat eosinophil-associated ribonuclease-1 and -2 and analysis of their biological activities. *Biochim. Biophys. Acta* 1638, 164–72.

Iyer, S., Holloway, D.E., Kumar, K., Shapiro, R., Acharya, K.R., 2005. Molecular recognition of human eosinophil-derived neurotoxin (RNase 2) by placental ribonuclease inhibitor. *J. Mol. Biol.* 347, 637–55.

Jackson, D.A., Symons, R.H., Berg, P., 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* 69, 2904–9.

Jardine, A.M., Leonidas, D.D., Jenkins, J.L., Park, C., Raines, R.T., Acharya, K.R., Shapiro, R., 2001. Cleavage of 3',5'-pyrophosphate-linked dinucleotides by ribonuclease A and angiogenin. *Biochemistry* 40, 10262–72.

Johnson, R.J., McCoy, J.G., Bingman, C.A., Phillips, G.N., Raines, R.T., 2007. Inhibition of human pancreatic ribonuclease by the human ribonuclease inhibitor protein. *J. Mol. Biol.* 368, 434–49.

Jones, W., 1920. The action of boiled pancreas extract on yeast nucleic acid. *Am. J. Physiol.* 52, 203–207.

Kabsch, W., 2010. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 125–32.

Kao, R.Y.T., Jenkins, J.L., Olson, K.A., Key, M.E., Fett, J.W., Shapiro, R., 2002. A small-molecule inhibitor of the ribonucleolytic activity of human angiogenin that possesses antitumor activity. *Proc. Natl. Acad. Sci. U. S. A.* 99, 10066–71.

Kartha, G., Bello, J., Harker, D., 1967. Tertiary structure of ribonuclease. *Nature* 213, 862–5.

Katoh, H., Yoshinaga, M., Yanagita, T., Ohgi, K., Irie, M., Beintema, J.J., Meinsma, D., 1986. Kinetic studies on turtle pancreatic ribonuclease: a comparative study of the base specificities of the B2 and P0 sites of bovine pancreatic ribonuclease A and turtle pancreatic ribonuclease. *Biochim. Biophys. Acta* 873, 367–71.

Kawanomoto, M., Motojima, K., Sasaki, M., Hattori, H., Goto, S., 1992. cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution of the mRNA. *Biochim. Biophys. Acta* 1129, 335–8.

Kazakou, K., 2008. *Structural studies on RNase A homologues.* Thesis (PhD). University of Bath, Bath.

Kazakou, K., Holloway, D.E., Prior, S.H., Subramanian, V., Acharya, K.R., 2008. Ribonuclease A homologues of the zebrafish: polymorphism, crystal structures of two representatives and their evolutionary implications. *J. Mol. Biol.* 380, 206–22.

Kendrew, J.C., 1963. Myoglobin and the structure of proteins. *Science* 139, 1259–66.

Kobe, B., Deisenhofer, J., 1993. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* 366, 751–6.

Kobe, B., Deisenhofer, J., 1994. The leucine-rich repeat: a versatile binding motif. *Trends Biochem. Sci.* 19, 415–21.

Kobe, B., Deisenhofer, J., 1995. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* 374, 183–6.

Kobe, B., Deisenhofer, J., 1996. Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. *J. Mol. Biol.* 264, 1028–43.

Kövér, K.E., Bruix, M., Santoro, J., Batta, G., Laurents, D. V, Rico, M., 2008. The solution structure and dynamics of human pancreatic ribonuclease determined by NMR spectroscopy provide insight into its remarkable biological activities and inhibition. *J. Mol. Biol.* 379, 953–65.

Kraft, N., Shortman, K., 1970. The phylogeny of the ribonuclease-ribonuclease inhibitor system: its distribution in tissues and its response during leukaemogenesis and aging. *Aust. J. Biol. Sci.* 23, 175–84.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., Karplus, K., 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* 77 Suppl 9, 114–22.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–97.

Kunitz, M., 1939. Isolation from beef pancreas of a crystalline protein possessing ribonuclease activity. *Science* 90, 112–3.

Kunitz, M., 1940. Crystalline ribonuclease. *J. Gen. Physiol.* 24, 15–32.

Ladner, J.E., Wladkowski, B.D., Svensson, L.A., Sjölin, L., Gilliland, G.L., 1997. X-ray structure of a ribonuclease A-uridine vanadate complex at 1.3 Å resolution. *Acta Crystallogr. D. Biol. Crystallogr.* 53, 290–301.

Landré, J.B.P., Hewett, P.W., Olivot, J.-M., Friedl, P., Ko, Y., Sachinidis, A., Moenner, M., 2002. Human endothelial cells selectively express large amounts of pancreatic-type ribonuclease (RNase 1). *J. Cell. Biochem.* 86, 540–52.

Lapteva, N., Nieda, M., Ando, Y., Nicol, a, Ide, K., Yamaura, a, Hatta-Ohashi, Y., Egawa, K., Juji, T., Tokunaga, K., 2001. Gene expression analysis in human monocytes, monocyte-derived dendritic cells, and alpha-galactosylceramide-pulsed monocyte-derived dendritic cells. *Biochem. Biophys. Res. Commun.* 289, 531–8.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–8.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.

LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., McCoy, J.M., 1993. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268, 23311–23317.

Lee, F.S., Auld, D.S., Vallee, B.L., 1989a. Tryptophan fluorescence as a probe of placental ribonuclease inhibitor binding to angiogenin. *Biochemistry* 28, 219–24.

Lee, F.S., Fox, E.A., Zhou, H.M., Strydom, D.J., Vallee, B.L., 1988. Primary structure of human placental ribonuclease inhibitor. *Biochemistry* 27, 8545–53.

Lee, F.S., Shapiro, R., Vallee, B.L., 1989b. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. *Biochemistry* 28, 225–30.

Lee, F.S., Vallee, B.L., 1989. Binding of placental ribonuclease inhibitor to the active site of angiogenin. *Biochemistry* 28, 3556–61.

Lehrer, R.I., Szklarek, D., Barton, A., Ganz, T., Hamann, K.J., Gleich, G.J., 1989. Antibacterial properties of eosinophil major basic protein and eosinophil cationic protein. *J. Immunol.* 142, 4428–34.

Leland, P.A., Schultz, L.W., Kim, B.M., Raines, R.T., 1998. Ribonuclease A variants with potent cytotoxic activity. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10407–12.

Leonidas, D.D., Boix, E., Prill, R., Suzuki, M., Turton, R., Minson, K., Swaminathan, G.J., Youle, R.J., Acharya, K.R., 2001. Mapping the ribonucleolytic active site of eosinophil-derived neurotoxin (EDN). High resolution crystal structures of EDN complexes with adenylic nucleotide inhibitors. *J. Biol. Chem.* 276, 15009–17.

Leonidas, D.D., Shapiro, R., Allen, S.C., Subbarao, G. V, Veluraja, K., Acharya, K.R., 1999a. Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth. *J. Mol. Biol.* 285, 1209–33.

Leonidas, D.D., Shapiro, R., Irons, L.I., Russo, N., Acharya, K.R., 1997. Crystal structures of ribonuclease A complexes with 5’-diphosphoadenosine 3'-phosphate and 5'-diphosphoadenosine 2'-phosphate at 1.7 Å resolution. *Biochemistry* 36, 5578–88.

Leonidas, D.D., Shapiro, R., Irons, L.I., Russo, N., Acharya, K.R., 1999b. Toward rational design of ribonuclease inhibitors: high-resolution crystal structure of a ribonuclease A complex with a potent 3',5'-pyrophosphate-linked dinucleotide inhibitor. *Biochemistry* 38, 10287–97.

Li, S., Sheng, J., Hu, J.K., Yu, W., Kishikawa, H., Hu, M.G., Shima, K., Wu, D., Xu, Z., Xin, W., Sims, K.B., Landers, J.E., Brown, R.H., Hu, G., 2013. Ribonuclease 4 protects neuron degeneration by promoting angiogenesis,

neurogenesis, and neuronal survival under stress. *Angiogenesis* 16, 387–404.

Liang, C.J., Yamashita, K., Kobata, a, 1980. Structural study of the carbohydrate moiety of bovine pancreatic ribonuclease B. *J. Biochem.* 88, 51–8.

Liao, Y.D., Huang, H.C., Leu, Y.J., Wei, C.W., Tang, P.C., Wang, S.C., 2000. Purification and cloning of cytotoxic ribonucleases from *Rana catesbeiana* (bullfrog). *Nucleic Acids Res.* 28, 4097–104.

Lindquist, R.N., Lynn, J.L., Lienhard, G.E., 1973. Possible transition-state analogs for ribonuclease. The complexes of uridine with oxovanadium(IV) ion and vanadium(V) ion. *J. Am. Chem. Soc.* 95, 8762–8.

Lipscomb, W.N., Coppola, J.C., Hartsuck, J. a., Ludwig, M.L., Muirhead, H., Searl, J., Steitz, T. a., 1966. The structure of carboxypeptidase A. *J. Mol. Biol.* 19, 423–441.

Lobban, P.E., Kaiser, a D., 1973. Enzymatic end-to end joining of DNA molecules. J. Mol. Biol. 78, 453–71.

Lomax, J.E., Bianchetti, C.M., Chang, A., Phillips, G.N., Fox, B.G., Raines, R.T., 2014. Functional evolution of ribonuclease inhibitor: insights from birds and reptiles. *J. Mol. Biol.* 426, 3041–56.

Luigetti, M., Lattante, S., Zollino, M., Conte, A., Marangi, G., Del Grande, A., Sabatelli, M., 2011. SOD1 G93D sporadic amyotrophic lateral sclerosis (SALS) patient with rapid progression and concomitant novel ANG variant. *Neurobiol. Aging* 32, 1924.e15–8.

Malhotra, A., 2009. Tagging for Protein Expression. *Methods Enzymol.* 463, 239–258.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.

McDonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238, 777–93.

McPherson, A., 2004. Introduction to protein crystallization. *Methods* 34, 254–65.

McPherson, A., Brayer, G., Cascio, D., Williams, R., 1986. The mechanism of binding of a polynucleotide chain to pancreatic ribonuclease. *Science* 232, 765–8.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., Lopez, R., 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* 41, W597–600.

Merrifield, B., 1986. Solid phase synthesis. *Science* 232, 341–7.

Messmore, J.M., Fuchs, D.N., Raines, R.T., 1995. Ribonuclease a: revealing structure-function relationships with semisynthesis. *J. Am. Chem. Soc.* 117, 8057–60.

Mizuta, K., Awazu, S., Yasuda, T., Kishi, K., 1990. Purification and characterization of three ribonucleases from human kidney: comparison with urine ribonucleases. *Arch. Biochem. Biophys.* 281, 144–51.

Moodie, S.L., Thornton, J.M., 1993. A study into the effects of protein binding on nucleotide conformation. *Nucleic Acids Res.* 21, 1369–1380.

Moore, S., Stein, W.H., 1973. Chemical structures of pancreatic ribonuclease and deoxyribonuclease. *Science* 180, 458–64.

Mosimann, S.C., Newton, D.L., Youle, R.J., James, M.N., 1996. X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 Å resolution. *J. Mol. Biol.* 260, 540–52.

Murdock, A.L., Grist, K.L., Hirs, C.H.W., 1966. On the dinitrophenylation of bovine pancreatic ribonuclease A kinetics of the reaction in water and 8M urea. *Arch. Biochem. Biophys.* 114, 375–390.

Murshudov, G.N., Skubák, P., Lebedev, A., Pannu, N.S., Steiner, R., Nicholls, R., Winn, M.D., Long, F., Vagin, A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* 67, 355–67.

Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 53, 240–255.

Navarro, S., Aleu, J., Jiménez, M., Boix, E., Cuchillo, C.M., Nogués, M. V, 2008. The cytotoxicity of eosinophil cationic protein/ribonuclease 3 on eukaryotic cell lines takes place through its aggregation on the cell membrane. *Cell. Mol. Life Sci.* 65, 324–37.

Nelson, D. L., Cox, M. M., 2008. *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Nitto, T., Dyer, K.D., Czapiga, M., Rosenberg, H.F., 2006. Evolution and function of leukocyte RNase A ribonucleases of the avian species, *Gallus gallus*. *J. Biol. Chem.* 281, 25622–34.

Nitto, T., Lin, C., Dyer, K.D., Wagner, R.A., Rosenberg, H.F., 2005. Characterization of a ribonuclease gene and encoded protein from the reptile, *Iguana iguana*. *Gene* 352, 36–44.

Nogués, M. V, Moussaoui, M., Boix, E., Vilanova, M., Ribó, M., Cuchillo, C.M., 1998. The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A. *Cell. Mol. Life Sci.* 54, 766–74.

Noguès, M. V, Vilanova, M., Cuchillo, C.M., 1995. Bovine pancreatic ribonuclease A as a model of an enzyme with multiple substrate binding sites. *Biochim. Biophys. Acta* 1253, 16–24.

Olson, K.A., Byers, H.R., Key, M.E., Fett, J.W., 2002. Inhibition of prostate carcinoma establishment and metastatic growth in mice by an antiangiogenin monoclonal antibody. *Int. J. Cancer* 98, 923–9.

Olson, K.A., Fett, J.W., French, T.C., Key, M.E., Vallee, B.L., 1995. Angiogenin antagonists prevent tumor growth in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 92, 442–6.

Olson, K.A., Verselis, S.J., Fett, J.W., 1998. Angiogenin is regulated in vivo as an acute phase protein. *Biochem. Biophys. Res. Commun.* 242, 480–3.

Oppenheim, J.J., Yang, D., 2005. Alarmins: chemotactic activators of immune responses. *Curr. Opin. Immunol.* 17, 359–65.

Panov, K.I., Kolbanovskaya, E.Y., Okorokov, A.L., Panova, T.B., Terwisscha van Scheltinga, A.C., Karpeisky MYa, Beintema, J.J., 1996. Ribonuclease A mutant His119 Asn: the role of histidine in catalysis. *FEBS Lett.* 398, 57–60.

Papageorgiou, A.C., Shapiro, R., Acharya, K.R., 1997. Molecular recognition of human angiogenin by placental ribonuclease inhibitor--an X-ray crystallographic study at 2.0 Å resolution. *EMBO J.* 16, 5162–77.

Parés, X., Llorens, R., Arús, C., Cuchillo, C.M., 1980. The reaction of bovine pancreatic ribonuclease A with 6-chloropurineriboside 5'-monophosphate. Evidence on the existence of a phosphate-binding sub-site. *Eur. J. Biochem.* 105, 571–9.

Parés, X., Noguès, M. V, de Llorens, R., Cuchillo, C.M., 1991. Structure and function of ribonuclease A binding subsites. *Essays Biochem.* 26, 89–103.

Parmenopoulou, V., Chatzileontiadou, D.S.M., Manta, S., Bougiatioti, S., Maragozidis, P., Gkaragkouni, D.-N., Kaffesaki, E., Kantsadi, A.L., Skamnaki, V.T., Zographos, S.E., Zounpoulakis, P., Balatsos, N.A.A., Komiotis, D., Leonidas, D.D., 2012. Triazole pyrimidine nucleosides as inhibitors of Ribonuclease A. Synthesis, biochemical, and structural evaluation. *Bioorg. Med. Chem.* 20, 7184–93.

Paubel, A., Violette, J., Amy, M., Praline, J., Meininger, V., Camu, W., Corcia, P., Andres, C.R., Vourc'h, P., 2008. Mutations of the ANG gene in French patients with sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* 65, 1333–6.

Penttinen, J., Pujianto, D.A., Sipila, P., Huhtaniemi, I., Poutanen, M., 2003. Discovery in silico and characterization in vitro of novel genes exclusively expressed in the mouse epididymis. *Mol. Endocrinol.* 17, 2138–2151.

Perutz, M.F., 1963. X-ray analysis of hemoglobin. *Science* 140, 863–9.

Piccoli, R., Olson, K.A., Vallee, B.L., Fett, J.W., 1998. Chimeric anti-angiogenin antibody cAb 26-2F inhibits the formation of human breast cancer xenografts in athymic mice. *Proc. Natl. Acad. Sci. U. S. A.* 95, 4579–83.

Pizzo, E., Buonanno, P., Di Maro, A., Ponticelli, S., De Falco, S., Quarto, N., Cubellis, M.V., D'Alessio, G., 2006. Ribonucleases and angiogenins from fish. *J. Biol. Chem.* 281, 27454–60.

Pizzo, E., Merlino, A., Turano, M., Russo Krauss, I., Coscia, F., Zanfardino, A., Varcamonti, M., Furia, A., Giancola, C., Mazzarella, L., Sica, F., D'Alessio, G., 2011. A new RNase sheds light on the RNase/angiogenin subfamily from zebrafish. *Biochem. J.* 433, 345–55.

Pizzo, E., Varcamonti, M., Di Maro, A., D Maro, A., Zanfardino, A., Giancola, C., D'Alessio, G., 2008. Ribonucleases with angiogenic and bactericidal activities from the Atlantic salmon. *FEBS J.* 275, 1283–95.

Plummer, T.H., 1968. Glycoproteins of bovine pancreatic juice. Isolation of ribonucleases C and D. *J. Biol. Chem.* 243, 5961–6.

Plummer, T.H., Hirs, C.H., 1964. On the structure of bovine pancreatic ribonuclease B. Isolation of a glycopeptide. *J. Biol. Chem.* 239, 2530–8.

Pollard, D.R., Nagyvary, J., 1973. Inhibition of pancreatic ribonuclease A by arabinonucleotides. *Biochemistry* 12, 1063–6.

Porath, J., Flodin, P., 1959. Gel Filtration: A Method for Desalting and Group Separation. *Nature* 183, 1657–1659.

Raines, R.T., 1998. Ribonuclease A. *Chem. Rev.* 98, 1045–1066.

Rayaprolu, S., Soto-Ortolaza, A., Rademakers, R., Uitti, R.J., Wszolek, Z.K., Ross, O.A., 2012. Angiogenin variation and Parkinson disease. *Ann. Neurol.* 71, 725–7; author reply 727–8.

Rees, W.A., Yager, T.D., Korte, J., von Hippel, P.H., 1993. Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* 32, 137–144.

Regnier, F.E., 1983. High-performance liquid chromatography of biopolymers. *Science* 222, 245–52.

Richards, F. M., Wyckoff, H. W., 1971. Bovine pancreatic ribonuclease. In: P. D. Boyer, ed. *The Enzymes.* New York: Academic Press. pp. 647-806.

Richards, F. M., Wyckoff, H. W., 1973. 1. Ribonuclease-S. In: D. C. Phillips, F. M. Richards, eds. *Atlas of Molecular Structures in Biology.* Oxford: Clarendon Press. pp. 1-54.

Roberts, G.C., Dennis, E.A., Meadows, D.H., Cohen, J.S., Jardetzky, O., 1969. The mechanism of action of ribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 62, 1151–8.

Rodgers, D.W., 1994. Cryocrystallography. *Structure* 2, 1135–40.

Rosenberg, H.F., Domachowske, J.B., 2001. Eosinophils, eosinophil ribonucleases, and their role in host defense against respiratory virus pathogens. *J. Leukoc. Biol.* 70, 691–8.

Rosenberg, H.F., Dyer, K.D., 1995. Human ribonuclease 4 (RNase 4): coding sequence, chromosomal localization and identification of two distinct transcripts in human somatic tissues. *Nucleic Acids Res.* 23, 4290–5.

Rosenberg, H.F., Dyer, K.D., 1996. Molecular cloning and characterization of a novel human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family. *Nucleic Acids Res.* 24, 3507–13.

Rossmann, M.G., Blow, D.M., 1962. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* 15, 24–31.

Roth, J.S., 1956. Ribonuclease. V. Studies on the properties and distribution of ribonuclease inhibitor in the rat. *Biochim. Biophys. Acta* 21, 34–43.

Roth, J.S., 1958a. Ribonuclease. VII. Partial purification and characterization of a ribonuclease inhibitor in rat liver supernatant fraction. *J. Biol. Chem.* 231, 1085–95.

Roth, J.S., 1958b. Ribonuclease. VIII. Studies on the inactive ribonuclease in the supernatant fraction of rat liver. *J. Biol. Chem.* 231, 1097–105.

Roth, J.S., 1962. Ribonuclease. IX. Further studies on ribonuclease inhibitor. *Biochim. Biophys. Acta* 61, 903–15.

Rudd, P.M., Scragg, I.G., Coghill, E., Dwek, R. a, 1992. Separation and analysis of the glycoform populations of ribonuclease B using capillary electrophoresis. *Glycoconj. J.* 9, 86–91.

Rudolph, B., Podschun, R., Sahly, H., Schubert, S., Schröder, J.M., Harder, J., 2006. Identification of RNase 8 as a novel human antimicrobial protein. *Antimicrob. Agents Chemother.* 50, 3194–6.

Rupp, B., 2010. *Biomolecular crystallography: principles, practice, and application to structural biology.* New York; Abingdon: Garland Science.

Russo, A., Acharya, K.R., Shapiro, R., 2001. Small molecule inhibitors of RNase A and related enzymes. *Methods Enzymol.* 341, 629–48.

Russo, N., Shapiro, R., 1999. Potent inhibition of mammalian ribonucleases by 3', 5'-pyrophosphate-linked nucleotides. *J. Biol. Chem.* 274, 14902–8.

Russo, N., Shapiro, R., Vallee, B.L., 1997. 5'-Diphosphoadenosine 3'-phosphate is a potent inhibitor of bovine pancreatic ribonuclease A. *Biochem. Biophys. Res. Commun.* 231, 671–4.

Schneider, R., Schneider-Scherzer, E., Thurnher, M., Auer, B., Schweiger, M., 1988. The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses a novel highly diversified protein superfamily with a common repetitive module. *EMBO J.* 7, 4151–6.

Schultz, L.W., Quirk, D.J., Raines, R.T., 1998. His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. *Biochemistry* 37, 8886–98.

Seilhean, D., Cazeneuve, C., Thuriès, V., Russaouen, O., Millecamps, S., Salachas, F., Meininger, V., Leguern, E., Duyckaerts, C., 2009. Accumulation of TDP-43 and alpha-actin in an amyotrophic lateral sclerosis patient with the K17I ANG mutation. *Acta Neuropathol.* 118, 561–73.

Seno, M., Futami, J., Tsushima, Y., Akutagawa, K., Kosaka, M., Tada, H., Yamada, H., 1995. Molecular cloning and expression of human ribonuclease 4 cDNA. *Biochim. Biophys. Acta* 1261, 424–6.

Shapiro, R., 2001. Cytoplasmic ribonuclease inhibitor. *Methods Enzymol.* 341, 611–28.

Shapiro, R., Fett, J.W., Strydom, D.J., Vallee, B.L., 1986. Isolation and characterization of a human colon carcinoma-secreted enzyme with pancreatic ribonuclease-like activity. *Biochemistry* 25, 7255–64.

Shapiro, R., Fox, E.A., Riordan, J.F., 1989. Role of lysines in human angiogenin: chemical modification and site-directed mutagenesis. *Biochemistry* 28, 1726–32.

Shapiro, R., Vallee, B.L., 1991. Interaction of human placental ribonuclease with placental ribonuclease inhibitor. *Biochemistry* 30, 2246–55.

Shapiro, R., Weremowicz, S., Riordan, J.F., Vallee, B.L., 1987. Ribonucleolytic activity of angiogenin: essential histidine, lysine, and arginine residues. *Proc. Natl. Acad. Sci. U. S. A.* 84, 8783–7.

She, J., Han, Z., Kim, T.-W., Wang, J., Cheng, W., Chang, J., Shi, S., Wang, J., Yang, M., Wang, Z.-Y., Chai, J., 2011. Structural insight into brassinosteroid perception by BRI1. *Nature* 474, 472–6.

Shortman, K., 1961. Studies on cellular inhibitors of ribonuclease I. The assay of the ribonuclease-inhibitor system, and the purification of the inhibitor from rat liver. *Biochim. Biophys. Acta* 51, 37–49.

Shortman, K., 1962. Studies on cellular inhibitors of ribonuclease. II. Some properties of the inhibitor from rat liver. *Biochim. Biophys. Acta* 55, 88–96.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539.

Sorrentino, S., 2010. The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett.* 584, 2194–200.

Sorrentino, S., Glitz, D.G., Hamann, K.J., Loegering, D.A., Checkel, J.L., Gleich, G.J., 1992. Eosinophil-derived neurotoxin and human liver ribonuclease. Identity of structure and linkage of neurotoxicity to nuclease activity. *J. Biol. Chem.* 267, 14859–65.

Sorrentino, S., Tucker, G.K., Glitz, D.G., 1988. Purification and characterization of a ribonuclease from human liver. *J. Biol. Chem.* 263, 16125–31.

Sowa, G.A., Hengge, A.C., Cleland, W.W., 1997. 18 O Isotope Effects Support a Concerted Mechanism for Ribonuclease A. *J. Am. Chem. Soc.* 119, 2319–2320.

Steidinger, T.U., Standaert, D.G., Yacoubian, T.A., 2011. A neuroprotective role for angiogenin in models of Parkinson's disease. *J. Neurochem.* 116, 334–41.

Strydom, D.J., Fett, J.W., Lobb, R.R., Alderman, E.M., Bethune, J.L., Riordan, J.F., Vallee, B.L., 1985. Amino acid sequence of human tumor derived angiogenin. *Biochemistry* 24, 5486–94.

Subramanian, V., Crabtree, B., Acharya, K.R., 2008. Human angiogenin is a neuroprotective factor and amyotrophic lateral sclerosis associated angiogenin variants affect neurite extension/pathfinding and survival of motor neurons. *Hum. Mol. Genet.* 17, 130–49.

Subramanian, V., Feng, Y., 2007. A new role for angiogenin in neurite growth and pathfinding: implications for amyotrophic lateral sclerosis. *Hum. Mol. Genet.* 16, 1445–53.

Suzuki, M., Saxena, S.K., Boix, E., Prill, R.J., Vasandani, V.M., Ladner, J.E., Sung, C., Youle, R.J., 1999. Engineering receptor-mediated cytotoxicity into human ribonucleases by steric blockade of inhibitor interaction. *Nat. Biotechnol.* 17, 265–70.

Swaminathan, G.J., Holloway, D.E., Veluraja, K., Acharya, K.R., 2002. Atomic resolution (0.98 Å) structure of eosinophil-derived neurotoxin. *Biochemistry* 41, 3341–52.

Tarragona-Fiol, A., Eggelte, H.J., Harbron, S., Sanchez, E., Taylorson, C.J., Ward, J.M., Rabin, B.R., 1993. Identification by site-directed mutagenesis of amino acids in the B2 subsite of bovine pancreatic ribonuclease A. *Protein Eng.* 6, 901–6.

Terzyan, S.S., Peracaula, R., de Llorens, R., Tsushima, Y., Yamada, H., Seno, M., Gomis-Rüth, F.X., Coll, M., 1999. The three-dimensional structure of human RNase 4, unliganded and complexed with d(Up), reveals the basis for its uridine selectivity. *J. Mol. Biol.* 285, 205–14.

Thiyagarajan, N., Ferguson, R., Subramanian, V., Acharya, K.R., 2012. Structural and molecular insights into the mechanism of action of human angiogenin-ALS variants in neurons. *Nat. Commun.* 3, 1121.

Thompson, J.E., Kutateladze, T.G., Schuster, M.C., Venegas, F.D., Messmore, J.M., Raines, R.T., 1995. Limits to Catalysis by Ribonuclease A. *Bioorg. Chem.* 23, 471–481.

Thompson, J.E., Raines, R.T., 1994. Value of general Acid-base catalysis to ribonuclease A. *J. Am. Chem. Soc.* 116, 5467–8.

Thompson, J.E., Venegas, F.D., Raines, R.T., 1994. Energetics of catalysis by ribonucleases: fate of the 2',3'-cyclic phosphodiester intermediate. *Biochemistry* 33, 7408–14.

Torrent, M., Badia, M., Moussaoui, M., Sanchez, D., Nogués, M.V., Boix, E., 2010. Comparison of human RNase 3 and RNase 7 bactericidal action at the Gram-negative and Gram-positive bacterial cell wall. *FEBS J.* 277, 1713–25.

Trautwein, K., Holliger, P., Stackhouse, J., Benner, S.A., 1991. Site-directed mutagenesis of bovine pancreatic ribonuclease: lysine-41 and aspartate-121. *FEBS Lett.* 281, 275–7.

Tsirkone, V.G., Dossi, K., Drakou, C., Zographos, S.E., Kontou, M., Leonidas, D.D., 2009. Inhibitor design for ribonuclease A: the binding of two 5'-phosphate uridine analogues. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 65, 671–7.

Usher, D.A., Erenrich, E.S., Eckstein, F., 1972. Geometry of the first step in the action of ribonuclease-A (in-line geometry/uridine2',3'-cyclic thiophosphate/^{31}P NMR). *Proc. Natl. Acad. Sci. U. S. A.* 69, 115–8.

Usher, D.A., Richardson, D.I., Eckstein, F., 1970. Absolute stereochemistry of the second step of ribonuclease action. *Nature* 228, 663–5.

Vaguine, a a, Richelle, J., Wodak, S.J., 1999. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D. Biol. Crystallogr.* 55, 191–205.

Van Es, M.A., Diekstra, F.P., Veldink, J.H., Baas, F., Bourque, P.R., Schelhaas, H.J., Strengman, E., Hennekam, E.A.M., Lindhout, D., Ophoff, R.A., van den Berg, L.H., 2009. A case of ALS-FTD in a large FALS pedigree with a K17I ANG mutation. *Neurology* 72, 287–8.

Van Es, M.A., Schelhaas, H.J., van Vught, P.W.J., Ticozzi, N., Andersen, P.M., Groen, E.J.N., Schulte, C., Blauw, H.M., Koppers, M., Diekstra, F.P., Fumoto, K., LeClerc, A.L., Keagle, P., Bloem, B.R., Scheffer, H., van Nuenen, B.F.L., van Blitterswijk, M., van Rheenen, W., Wills, A.-M., Lowe, P.P., Hu, G., Yu, W., Kishikawa, H., Wu, D., Folkerth, R.D., Mariani, C., Goldwurm, S., Pezzoli, G., Van Damme, P., Lemmens, R., Dahlberg, C., Birve, A., Fernández-Santiago, R., Waibel, S., Klein, C., Weber, M., van der Kooi, A.J., de Visser, M., Verbaan, D., van Hilten, J.J., Heutink, P., Hennekam, E.A.M., Cuppen, E., Berg, D., Brown, R.H.,

Silani, V., Gasser, T., Ludolph, A.C., Robberecht, W., Ophoff, R.A., Veldink, J.H., Pasterkamp, R.J., de Bakker, P.I.W., Landers, J.E., van de Warrenburg, B.P., van den Berg, L.H., 2011. Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. *Ann. Neurol.* 70, 964–73.

Vera, L., Stura, E.A., 2014. Strategies for Protein Cryocrystallography. *Cryst. Growth Des.* 14, 427–435.

Vescia, S., Tramontano, D., Augusti-Tocco, G., D'Alessio, G., 1980. In vitro studies on selective inhibition of tumor cell growth by seminal ribonuclease. *Cancer Res.* 40, 3740–4.

Vicentini, a M., Hemmings, B. a, Hofsteenge, J., 1994. Residues 36-42 of liver RNase PL3 contribute to its uridine-preferring substrate specificity. Cloning of the cDNA and site-directed mutagenesis studies. *Protein Sci.* 3, 459–66.

Vicentini, a M., Kote-Jarai, Z., Hofsteenge, J., 1996. Structural determinants of the uridine-preferring specificity of RNase PL3. *Biochemistry* 35, 9128–32.

Vicentini, A.M., Kieffer, B., Matthies, R., Meyhack, B., Hemmings, B.A., Stone, S.R., Hofsteenge, J., 1990. Protein chemical and kinetic characterization of recombinant porcine ribonuclease inhibitor expressed in Saccharomyces cerevisiae. *Biochemistry* 29, 8827–34.

Wang, J.H., 1968. Facilitated proton transfer in enzyme catalysis. It may have a crucial role in determining the efficiency and specificity of enzymes. *Science* 161, 328–34.

Weil, L., Seibles, T.S., 1955. Photooxidation of crystalline ribonuclease in the presence of methylene blue. *Arch. Biochem. Biophys.* 54, 368–77.

Wilcheck, M., Chaiken, I., 2000. An overview of affinity chromatography. In: P. Bailon, G. K. Ehrlich, W. J. Fung, W. Berthold, eds. *Affinity*

Chromatography: methods and protocols. Totowa, NJ: Humana Press. pp. 1-6.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E. a, Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* 67, 235–42.

Winter, G., Lobley, C.M.C., Prince, S.M., 2013. Decision making in xia2. *Acta Crystallogr. D. Biol. Crystallogr.* 69, 1260–73.

Witzel, H., 1963. The function of the pyrimidine base in the ribonuclease reaction. *Prog. Nucleic Acid Res. Mol. Biol.* 2, 221–258.

Wlodawer, a, Miller, M., Sjölin, L., 1983. Active site of RNase: neutron diffraction study of a complex with uridine vanadate, a transition-state analog. *Proc. Natl. Acad. Sci. U. S. A.* 80, 3628–31.

Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M., 2008. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* 275, 1-21.

Wu, D., Yu, W., Kishikawa, H., Folkerth, R.D., Iafrate, A.J., Shen, Y., Xin, W., Sims, K., Hu, G.-F., 2007. Angiogenin loss-of-function mutations in amyotrophic lateral sclerosis. *Ann. Neurol.* 62, 609–17.

Wu, Y., Mikulski, S.M., Ardelt, W., Rybak, S.M., Youle, R.J., 1993. A cytotoxic ribonuclease. Study of the mechanism of onconase cytotoxicity. *J. Biol. Chem.* 268, 10686–93.

Yang, D., Chen, Q., Rosenberg, H.F., Rybak, S.M., Newton, D.L., Wang, Z.Y., Fu, Q., Tchernev, V.T., Wang, M., Schweitzer, B., Kingsmore, S.F., Patel, D.D., Oppenheim, J.J., Howard, O.M.Z., 2004. Human ribonuclease A superfamily members, eosinophil-derived neurotoxin and pancreatic

ribonuclease, induce dendritic cell maturation and activation. *J. Immunol.* 173, 6134–42.

Yang, D., Rosenberg, H.F., Chen, Q., Dyer, K.D., Kurosaka, K., Oppenheim, J.J., 2003. Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with chemotactic activities for dendritic cells. *Blood* 102, 3396–403.

Yasuda, T., Mizuta, K., Sato, W., Kishi, K., 1990. Purification and characterization of a ribonuclease from human spleen. Immunological and enzymological comparison with nonsecretory ribonuclease from human urine. *Eur. J. Biochem.* 191, 523–9.

Yasuda, T., Nadano, D., Takeshita, H., Kishi, K., 1993. Two distinct secretory ribonucleases from human cerebrum: purification, characterization and relationships to other ribonucleases. *Biochem. J.* 296 (Pt 3), 617–25.

Yazdanbakhsh, M., Tai, P.C., Spry, C.J., Gleich, G.J., Roos, D., 1987. Synergism between eosinophil cationic protein and oxygen metabolites in killing of schistosomula of *Schistosoma mansoni*. *J. Immunol.* 138, 3443–7.

Yuan, A., Rao, M. V., Veeranna, Nixon, R.A., 2012. Neurofilaments at a glance. *J. Cell Sci.* 125, 3257–3263.

Zelenko, O., Neumann, U., Brill, W., Pieles, U., Moser, H.E., Hofsteenge, J., 1994. A novel fluorogenic substrate for ribonucleases. Synthesis and enzymatic characterization. *Nucleic Acids Res.* 22, 2731–9.

Zhang, J., Dyer, K.D., Rosenberg, H.F., 2002. RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta. *Nucleic Acids Res.* 30, 1169–75.

Zhang, J., Dyer, K.D., Rosenberg, H.F., 2003. Human RNase 7: a new cationic ribonuclease of the RNase A superfamily. *Nucleic Acids Res.* 31, 602–7.

Zhao, W., Beintema, J.J., Hofsteenge, J., 1994. The amino acid sequence of iguana (*Iguana iguana*) pancreatic ribonuclease. *Eur. J. Biochem.* 219, 641–6.

Zhao, W., Kote-Jarai, Z., van Santen, Y., Hofsteenge, J., Beintema, J.J., 1998. Ribonucleases from rat and bovine liver: purification, specificity and structural characterization. *Biochim. Biophys. Acta* 1384, 55–65.

Zhou, H.M., Strydom, D.J., 1993. The amino acid sequence of human ribonuclease 4, a highly conserved ribonuclease that cleaves specifically on the 3' side of uridine. *Eur. J. Biochem.* 217, 401–10.

Zou, Z.-Y., Wang, X.-N., Liu, M.-S., Sun, Q., Li, X.-G., Cui, L.-Y., Kong, J., 2012. Identification of a novel missense mutation in angiogenin in a Chinese amyotrophic lateral sclerosis cohort. *Amyotroph. Lateral Scler.* 13, 270–5.

Appendix I – pET-46 Ek/LIC vector map

The Novagen pET-46 Ek/LIC vector is a bacterial vector for ligation-independent cloning (LIC) to express fusion proteins with an N-terminal cleavable 6×His-tag, and an optional C-terminal S-tag.

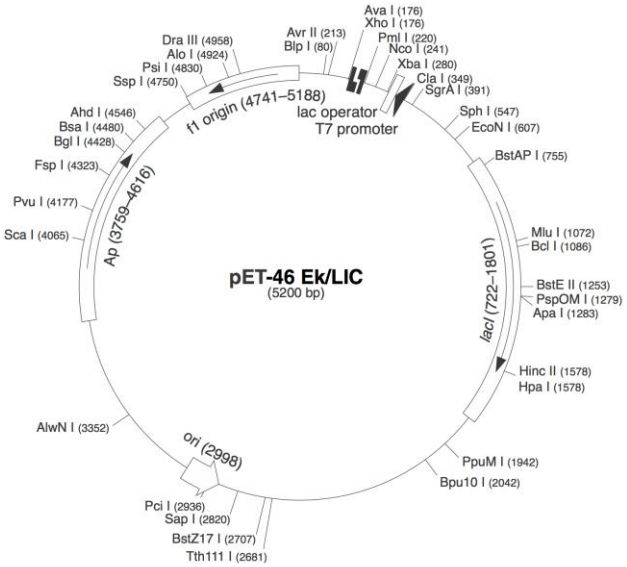
Novagen®

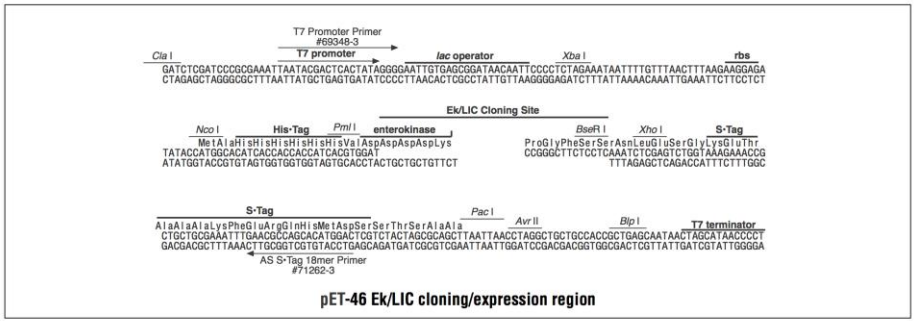
pET-46 Ek/LIC Vector

TB397 0903

	Cat. No.
pET-46 Ek/LIC Cloning Kit	71335-3
pET-46 Ek/LIC sequence landmarks	
T7 promoter	315–331
T7 transcription start	314
His•Tag® coding sequence	221–238
Ek/LIC cloning site	186–214
S•Tag™ coding sequence	125–169
T7 terminator	26–73
lacI coding sequence	722–1801
pBR322 origin	2998
bla (Ap^R) coding sequence	3759–4616
f1 origin	4741–5188

The pET-46 Ek/LIC vector is prepared for rapid, directional cloning of PCR-amplified DNA for high-level expression of polypeptides. Using specifically designed primers for amplification and the pET-46 Ek/LIC Cloning Kit (Cat. No. 71335-3), inserts can be efficiently cloned without the need for restriction digestion or ligation. Fusion proteins contain an N-terminal cleavable His•Tag® sequence. Fusion to an optional C-terminal S•Tag™ sequence can also be created for detection, purification and quantification of fusion proteins. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circle map. The f1 origin is oriented so that infection with helper phage will produce virions containing single stranded DNA that corresponds to the coding strand. Therefore, single stranded sequencing should be performed using the T7 Terminator Primer (Cat. No. 69337-3).





Appendix II – Optimised expression and purification protocols of wild type PL_3 from $BL21C^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 strain

Construct $BL21C^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 was prepared in this study and used to express and purify wild type PL_3. TB (Terrific Broth) media supplemented with 100 μg/ml ampicillin was inoculated with 1% overnight culture of $BL21C^+$(DE3)-RIPL-pET-46 Ek/LIC-PL_3 cells, and incubated at 37 ℃ with shaking at 200 rpm. Protein expression was induced by the addition of 1 mM IPTG to cells in the mid-exponential phase of growth (indicated by $OD_{600} \geq 0.6$). The induced cells were further incubated at 16 ℃ with shaking at 200 rpm for 16 hours, and harvested by centrifugation at 15, 000 g (Beckman JLA 10.500 rotor) for 10 minutes.

Cells from 0.5 L bacterial culture were monodispersed in 50 ml lysis buffer (50 mM Tris-HCl, pH 8.0) and ruptured using the TS Series Benchtop cell disruptor (Constant Systems Ltd., Daventry, Northants, UK) at 24,000 psi. The lysate was centrifuged at 39,000 g (Beckman JLA 25.50 rotor) for 20 minutes to harvest inclusion bodies.

The inclusion body pellet was solubilised in 10 ml solubilisation buffer (7 M guanidine hydrochloride, 0.15 M reduced glutathione, 0.1 M Tris-HCl pH 8.0, 2 mM EDTA) for 2 hours. The solubilised protein was separated from the solid matter by centrifugation at 70, 000 g for 30 minutes (Beckman JLA 25.50 rotor), and added dropwisely into 500 ml refolding buffer (0.5 M L-arginine, 0.6 mM oxidized glutathione) with gentle stirring. After the addition of the protein sample, the refolding solution was left undisturbed at 4 ℃ overnight.

The PL_3 protein was first purified by cation exchange chromatography. A self-packed column with SP-Sepharose Fast Flow beads (GE Healthcare, UK) was equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl). Precipitated protein was removed from the refolding solution by centrifugation at 15, 000 g for 30 minutes (Beckman JLA 10.500 rotor); solubilised protein sample was diluted to 2.5 L with Milli-Q H_2O and loaded onto the column. Subsequently the column was re-equilibrated with equilibration buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl). PL_3 was eluted with an ascending NaCl gradient in equilibration buffer (from 0.1 M to 1.5 M NaCl). Fractions with raised UV trace were analysed by gel electrophoresis.

The protein sample was further purified using reverse phase HPLC chromatography. A Jupiter C4 column (Phenomenex, USA) was used. The column was washed with 95% buffer B (HPLC-grade acetonitrile with 0.08% trifluoroacetic acid (TFA)) and equilibrated with 95% buffer A (degassed HPLC-grade H_2O with 0.1% TFA) prior to protein loading. The column was re-equilibrated with buffer A after the protein was loaded; and the protein was eluted with an increasing Buffer B concentration. Fractions with raised UV trace were pooled and lyophilised overnight. The lyophilised protein sample was dissolved in HPLC-grade H_2O, and analysed by gel electrophoresis.